



# ESCOLA NAVAL

ta sainte & bi-faire



Joana Alves de Melo Jerónimo

**Análise Comparativa entre diferentes fontes de  
dados oceanográficos para a região da  
Macaronésia**

**Dissertação para obtenção do Grau de Mestre em  
Ciências Militares Navais, na especialidade de Marinha**



Alfeite

2021





# ESCOLA NAVAL

talant de bi-faire



**Joana Alves de Melo Jerónimo**

*Análise Comparativa entre diferentes fontes de dados  
oceanográficos para a região da Macaronésia*

Dissertação para obtenção do Grau de Mestre em  
Ciências Militares Navais, na especialidade de Marinha

**Orientação de:** Professora Doutora Luísa Lamas

**Co-orientação de:** Professor Doutor Ricardo Pinto Moura

*O Aluno Mestrando,*

*O Orientador,*

---

Joana Jerónimo

---

Luísa Lamas

Alfeite  
2021



"The important thing is not to stop questioning.

Curiosity has its own reason for existing."

Albert Einstein



# Agradecimentos

À Professora Doutora Luísa Lamas, o meu especial agradecimento por ter-me ajudado a encontrar um tema que me motivasse e que fornecesse contributos para a comunidade científica de oceanografia e para a Marinha Portuguesa. Agradeço a sua disponibilidade durante todo o desenvolvimento deste trabalho. Apelou ao meu espírito crítico e autonomia sem nunca deixar de partilhar o seu conhecimento e experiência.

Ao Professor Doutor Ricardo Moura, que me deu o seu contributo e concedeu uma especial ajuda na análise dos dados.

Ao Comandante Santos Fernandes e Doutora Rita Esteves do Instituto Hidrográfico por me terem ajudado a criar o tema inicial que me cativou e me levou a escolher um tema na área da oceanografia.

A toda a minha família. Ao meu pai por me ter ensinado a resiliência e que só com trabalho árduo se alcança o sucesso. À minha mãe por ser um exemplo constante de força e de trabalho. À minha irmã por ter sempre sido um apoio e um exemplo a seguir. E ao Tiago, por ser o meu porto de abrigo. Agradeço a todos a força e os valores que sempre me incutiram e que fazem de mim o que sou.

Ao meu curso Capitão-tenente Raúl Alexandre Cascais, em especial às minhas camaradas ASPOF Ana Sintra, CAD Susana Bernardo e ASPOF Teresa Vidiinha por uma vez mais terem estado ao meu lado durante esta nova etapa.

Por último, gostaria de expressar os meus agradecimentos a todos aqueles que contribuíram para a minha formação e, em especial, para a realização deste trabalho.





# Resumo

Este estudo foca-se na análise dos registos de Temperatura de Superfície do Mar (TSM) oriundos de boias derivantes (GDP), de um modelo numérico (MOD) e de satélites (SAT), na região da Macaronésia, no período de 30 de junho de 2019 a 30 de junho de 2020. A principal finalidade é identificar as limitações e potencialidades do uso de cada um destes métodos, nesta área.

Para tal, foram utilizados o *Root Mean Square Error* (RMSE), *Mean Bias Error* (MBE), *Mean Absolute Error* (MAE) e correlação de Kendall para avaliar a diferença entre os três conjuntos de dados. Comparou-se estes erros por subárea de forma a estudar a variabilidade espaciotemporal. Dadas as características de cada metodologia, considerou-se as  $TSM_{GDP}$  como a *ground truth*.

As boias medem a TSM através de termístores e apenas têm uma cobertura de 77%. Estas medem os dados *in-situ* (mais confiável) e em tempo real. Quanto ao MOD, este assimila dados de boias fundeadas, derivantes, navios e imagens satélite, e nunca poderá ser independente destes. No entanto, fornece dados instantâneos a todas as horas do dia, efetua previsões e análises aos seus registos e tem uma abrangência global espacial e temporal. Por último, o SAT combina vários resultados de imagens microondas e infravermelhos. Carecem de calibração, validação por dados *in-situ* e tem uma cobertura global. O resultado das diferenças é a medição de diferentes profundidades da camada de superfície.

Neste estudo, as diferenças entre os conjuntos de dados verificam-se maiores perante fenómenos de mesoescala, nomeadamente *eddies*, e estes encontram-se mais frequentes nas regiões do Arquipélago dos Açores e Cabo Verde, nos meses mais frios.

**Palavras-chave:** Comparação, TSM, GDP, satélites, modelo numérico, Macaronésia



# Abstract

This study focuses on the analysis of Sea Surface Temperature (SST) records from drifting buoys, a numerical model (MOD) and satellites (SAT), in the Macaronesia region, from June 30, 2019 to June 30, 2020. The main purpose is to identify limitations and the capabilities of the usage of each of these in this particular area.

The *Root Mean Square Error* (RMSE), *Mean Bias Error* (MBE), *Mean Absolute Error* (MAE) and Kendall correlation were used to evaluate the differences between these three datasets. These errors were compared by sub-area in order to study the spatiotemporal variability. Regarding the characteristics of each methods, the GDP data was considered the study's ground truth.

Buoys use thermistor to take measurements of SST but only have a 77% coverage. They measure data in-situ (most reliable) and in real time. As for the MOD, it assimilates data from anchored buoys, drifts, ships and satellite images, and can never be independent from them. Moreover, it offers instantaneous data at each hour of day, it provides forecast and analysis of its records, and has a global spatial and temporal scope. SAT combines various microwave and infrared imaging results. Requires calibration and validation by data in-situ. The result of the differences is the difference in surface layer depths.

In this study, the differences between the data sets are greater in the face of mesoscale phenomena (meanders and eddies), and these are more frequent in the regions of the Archipelago of the Azores and Cape Verde in the colder months.

**Keywords:** Comparison, SST, GDP, Satellite, Numerical Model, Macaronesia



# Índice

<b>Introdução</b>	<b>1</b>
<b>1 Dados e Métodos</b>	<b>13</b>
1.1 Descrição dos dados . . . . .	13
1.1.1 Dados das boias derivantes do GDP . . . . .	13
1.1.2 Dados do modelo numérico . . . . .	17
1.1.3 Dados dos Satélites . . . . .	18
1.2 Métodos estatísticos aplicados . . . . .	19
1.3 Tratamento dos Dados . . . . .	21
<b>2 Análise dos Resultados</b>	<b>27</b>
2.1 Análise da TSM na Macaronésia através dos dados GDP . . . . .	27
2.2 Análise Comparativa da TSM . . . . .	31
2.3 Análise das boias lançadas pelo NRP Sagres . . . . .	37
<b>3 Discussão dos Resultados</b>	<b>55</b>
3.1 Boias lançadas pelo NRP Sagres . . . . .	58
<b>Conclusão</b>	<b>58</b>
<b>Bibliografia</b>	<b>63</b>
<b>Apêndices</b>	<b>67</b>
<b>A Processo de obtenção do ficheiro dos dados sujeitos a análise</b>	<b>67</b>
A.1 Código de criação do Excel . . . . .	68
<b>B Descrição do tratamento dos dados</b>	<b>73</b>
<b>C Variação temporal dos registos TSM</b>	<b>81</b>
<b>D Erros MAE, MBE e RMSE das subáreas</b>	<b>85</b>



# Lista de Figuras

1	Perfis de temperatura da camada próxima à superfície (aproximadamente até 10 m de profundidade) do oceano - (a) condições diurnas e noturnas com ventos fortes; (b) condições diurnas com ventos fracos e elevada intensidade de radiação solar (retirada de Donlon et al., 2001).	3
2	Mapa geográfico da área de investigação. . . . .	6
3	Esboço da circulação geral próxima à superfície do Oceano Atlântico Norte (retirado de Barton, 2001). . . . .	10
4	Volume de transporte total do campo médio de densidade de longo prazo com assunção geostrófica e até 200m de profundidade(A) e vetores de corrente geostrófica calculado a partir do campo de densidade média de longo prazo e correntes médias observadas (B) esperados para a região em estudo (retirado de Barton, 2001). . . . .	10
1.1	Esboço do funcionamento das diferentes partes uma boia derivante (retirada de NOAA, s.d.-c). . . . .	14
1.2	Esquema de uma boia derivante SVP(retirado do <a href="https://rsaqua.co.uk/product/svp/">https://rsaqua.co.uk/product/svp/</a> ). . . . .	16
1.3	Mapa de densidade de cobertura de boias de janeiro de 2021 (Retirado de NOAA, s.d.-c). . . . .	17
1.4	Mapa de densidade de cobertura de boias de junho de 2021 (Retirado de NOAA, s.d.-c). . . . .	17
1.5	Variação temporal do número de boias a registar diariamente, em simultâneo, na área de estudo. . . . .	23
1.6	<i>Boxplot</i> representativo do número de boias GDP a registar diariamente na área de estudo. . . . .	24
1.7	Posições geográficas das boias GDP na AI, entre 30 de junho de 2019 e 30 de junho de 2020. . . . .	24
1.8	Posições geográficas das boias GDP nas subáreas, identificadas com as respetivas cores, entre 30 de junho de 2019 e 30 de junho de 2020.	25
1.9	Quantidade de registos (Em cima) e boias (Em baixo) por subárea.	26

2.1	Variação global temporal dos registos $TSM_{GDP}$ . . . . .	28
2.2	Posições geográficas das boias GDP entre 1 julho 2019 (losango verde) e 31 de julho 2019 (losango vermelho). . . . .	29
2.3	Posições geográficas das boias GDP entre 1 de dezembro 2019 (losango verde) e 31 de dezembro 2019 (losango vermelho). . . . .	30
2.4	Posições geográficas das boias GDP entre 1 de maio 2019 (losango verde) e 31 de maio 2019 (losango vermelho). . . . .	31
2.5	Matriz de correlação de Kendall entre dados TSM. . . . .	33
2.6	Valores dos MBE e número de registos por cada subárea definida no capítulo 1. . . . .	34
2.7	Valores dos MAE e número de registos por cada subárea definida no capítulo 1. . . . .	34
2.8	Valores dos RMSE e número de registos por cada subárea definida no capítulo 1. . . . .	35
2.9	Histograma das diferenças entre TSMGDP e TSMMOD (Azul) e TSMGDP e TSMSAT (Vermelho) em percentagem. . . . .	36
2.10	Correlação entre MAE e distância entre posições das boias durante o registo e a posição do registo MOD (Azul) ou SAT (vermelho) na malha. . . . .	36
2.11	Matriz de correlação entre os valores de TSM e distância entre posições das boias durante o registo e a posição do registo MOD e SAT na malha. . . . .	37
2.12	Posições geográficas das boias lançadas pelo NRP Sagres na AI durante o período de estudo. . . . .	38
2.13	Posições geográficas da boia 1301592, entre 9 de janeiro 2020 e 29 de junho 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identifica-se com a letra A, uma situação onde os valores previstos da TSMMOD e TSMSAT registaram maiores diferenças em relação ao TSMGDP. . . . .	39
2.14	Variação temporal do BE, AE e distância entre a posição geográfica da boia e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD. . . . .	39
2.15	Variação temporal da TSM (azul: TSMGDP, laranja: TSMMOD; verde: TSMSAT) correspondente ao trajeto da boia 1301592 durante o período analisado (09 de janeiro a 29 de junho 2020). O círculo a cinzento claro identifica o período correspondente à situação A. . .	40



2.16	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301592 durante situação A. . . . .	41
2.17	Posições geográficas da boia 6202656, entre 9 de janeiro 2020 e 29 de junho 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identifica-se com a letra B, uma situação onde os valores previstos da TSMMOD e TSMSAT registaram maiores diferenças em relação ao TSMGDP. . . . .	41
2.18	Variação temporal do BE, AE e distância entre a posição geográfica da boia 6202656 e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD, entre 9 de janeiro 2020 e 29 de junho 2020. . . . .	42
2.19	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 6202656 durante o período analisado (09 de janeiro a 29 de junho 2020). O círculo a cinzento claro identifica o período correspondente à situação B. . . . .	43
2.20	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 6202656 durante situação B. . . . .	43
2.21	Posições geográficas da boia 6202657, entre 9 de janeiro 2020 e 29 de junho 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identifica-se com a letra C, uma situação onde os valores previstos da TSMMOD e TSMSAT registaram maiores diferenças em relação ao TSMGDP. . . . .	44
2.22	Variação temporal do BE, AE e distância entre a posição geográfica da boia 6202657 e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD. . . . .	45
2.23	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 6202657 durante o período analisado (1 a 15 de janeiro 2020). O círculo a cinzento claro identifica o período correspondente à situação C. . . . .	45
2.24	Variação temporal da TSM preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 6202657 durante a situação C. . . . .	46

2.25	Posições geográficas da boia 1301589, entre 1 e 15 janeiro 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identificam-se com duas letras, E e D, situações onde os valores previstos da TSMMOD e TSMSAT registaram maiores diferenças em relação ao TSMGDP. . . . .	47
2.26	Variação temporal do BE, AE e distância entre a posição geográfica da boia 1301589 e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD. . . . .	48
2.27	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301589 durante o período analisado. A sombreado identifica-se o período correspondente às situações D e E. . . . .	48
2.28	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 61301589 durante a situação D. . . . .	49
2.29	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 61301589 durante a situação E. . . . .	50
2.30	Posições geográficas da boia 1301590, entre 17 de janeiro e 27 de junho de 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identifica-se com as letras F e G, situações onde os valores previstos da TSMMOD e TSMSAT registaram maiores diferenças em relação ao TSMGDP. . . . .	50
2.31	Variação temporal do BE, AE e distância entre a posição geográfica da boia 1301590 e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD. . . . .	51
2.32	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301590 durante o período analisado (17 de janeiro a 27 de junho 2020). O círculo a cinzento claro identifica o período correspondente à situação F e o cinzento escuro identifica o correspondente à situação G. . . . .	52
2.33	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301590 durante a situação F. . . . .	53
2.34	Variação temporal da TSM (preto: $TSM_{GDP}$ , azul: $TSM_{MOD}$ ; vermelho: $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301590 durante a situação G. . . . .	53

A.1	Serie temporal dos três conjuntos de dados em que a primeira coluna corresponde aos dados MOD, a segunda aos dados SAT, e a terceira aos GDP. . . . .	67
A.2	Imagem retirada do ficheiro Excel data_merge.xlsx . . . . .	72
B.1	Gráficos probabilísticos de quantis. . . . .	73
B.2	Gráfico de dispersão com as respetivas retas de regressão a verde e retas $x=y$ a tracejado a vermelho, sendo o gráfico da esquerda relativo aos dados MOD e o da direita aos dados SAT. . . . .	74
B.3	Gráfico de dispersão dos resíduos MOD (em cima) e SAT (em baixo) com os valores superiores a $1^{\circ}\text{C}$ a vermelho. . . . .	75
B.4	Histogramas da frequência de diferenças superiores a $1^{\circ}\text{C}$ entre dados MOD(A) e SAT(B) em relação aos GDP . . . . .	76
B.5	Posições geográficas da boia 4101718, entre 2020. A primeira e última posição estão marcadas por um círculo laranja. . . . .	77
B.6	Variação temporal de dados da boia 4101718 de TSM GDP, SAT e MOD. . . . .	77
B.7	Posições geográficas da boia 4201517, entre 22 de outubro a 2 de maio de 2020. A primeira e última posição estão marcadas por um círculo laranja. . . . .	78
B.8	Variação temporal de dados da boia 4201517 de TSM GDP, SAT e MOD . . . . .	78
C.1	Variação temporal dos registos $TSM_{GDP}$ por cada uma das subáreas definidas no Capítulo 1. . . . .	83
C.2	Variação do MAE por mês de estudo . . . . .	83
C.3	Variação do RMSE por mês de estudo . . . . .	84



# Lista de Tabelas

0.1	Limites geográficos da área de interesse. . . . .	6
2.1	Resultados da análise estatística da diferença entre os valores de $TSM_{GDP}$ e cada um dos valores de $TSM_{MOD}$ e $TSM_{SAT}$ . . . . .	32
2.2	Valores de $\sigma$ , RMSE, MAE e MBE da amostra em estudo. . . . .	33
3.1	Tabela síntese dos erros das boias lançadas pelo NRP Sagres . . . . .	58
B.1	Tabela resumo dos elementos estatísticos RMSE, MAE, MBE das boias 4201517, 4101718, dos registos globais e dos registos globais após a remoção dos registos da boia 4101718. . . . .	79
C.1	Número de registos e número de boias por subárea geográfica, para o período entre 1 de julho 2019 e 30 de junho 2020. . . . .	81
D.1	Registos dos erros MAE, MBE e RMSE de cada uma das subáreas definidas no Capítulo 1, e sua média do conjunto de dados. . . . .	86



# Lista de Equações

1.1	Equação do desvio padrão . . . . .	20
1.2	Equação do MBE . . . . .	20
1.3	Equação do MAE . . . . .	20
1.4	Equação do RMSE . . . . .	20
1.5	Fórmula do coeficiente de correlação de Kendall . . . . .	21





# Lista de Siglas e Acrónimos

AE	<i>Absolute Error.</i>
AI	<i>Área de Interesse.</i>
AVHRR	<i>Advanced Very High Resolution Radiometer.</i>
AzC	<i>Azores Current.</i>
CaC	<i>Canary Current.</i>
CMEMS	<i>Copernicus Marine Environment Monitoring Service.</i>
GDP	<i>Global Drifter Program.</i>
GHRSSST	<i>Group for High Resolution Sea Surface Temperature.</i>
GMPE	<i>GHRSSST Multi-Product Ensemble.</i>
GTS	<i>Global Telecommunications System.</i>
HN	<i>Hemisfério Norte.</i>
IH	<i>Instituto Hidrográfico.</i>
MAE	<i>Mean Absolute Error.</i>
MBE	<i>Mean Bias Error.</i>
MOD	<i>Modelo numérico em estudo.</i>
NAC	<i>North Atlantic Current.</i>
NEC	<i>North Equatorial Current.</i>
NOAA	<i>National Oceanographic and Atmospheric Administration.</i>
NRP	<i>Navio da República Portuguesa.</i>
PC	<i>Portugal Current.</i>

RMSE	<i>Root Mean Square Error.</i>
SAT	Produto de imagens satélite em estudo.
SVP	<i>Surface Velocity Program.</i>
$TSM_{GDP}$	dados de temperatura de superfície do mar obtidos pelas boias derivantes do GDP.
$TSM_{MOD}$	dados de temperatura de superfície do mar obtidos pelo modelo numérico em estudo.
$TSM_{SAT}$	dados de temperatura de superfície do mar obtidos pelas medições efetuadas por satélites.
TSM	Temperatura de Superfície do Mar.
VIIRS	<i>Visible Infrared Imaging Radiometer Suite.</i>
WMO	<i>World Meteorological Organization.</i>
ZCIT	<i>Zona de Convergência Intertropical.</i>

# Introdução

A Temperatura de Superfície do Mar (TSM) é um parâmetro ambiental fundamental no estudo da meteorologia e climatologia, e é considerada uma das variáveis climáticas essenciais pela Organização Meteorológica Mundial (Alvera-Azcárate et al., 2011). O estudo da sua variabilidade espaciotemporal é essencial para o conhecimento dos processos que ocorrem na camada de superfície do mar (como, por exemplo, os fluxos de energia entre o sistema atmosfera-oceano) e por conseguinte, a deteção das suas alterações, que podem ser de pequena ou grande escala, beneficiando várias das suas aplicações operacionais, como modelos de previsão meteorológica (Alvera-Azcárate et al., 2011; Donlon et al., 2001).

A atmosfera é transparente a ondas eletromagnéticas com comprimentos de onda entre 8 e 13  $\mu\text{m}$ . Como tal, os sensores remotos são configurados para determinados intervalos de janelas atmosféricas neste espectro de modo que, com base na Lei de Planck a partir da qual se deduz a Lei de deslocamento de *Wien*: “o comprimento de onda da energia eletromagnética irradiada de um corpo negro é inversamente proporcional à temperatura do mesmo”, se obtenha os valores de temperatura da Terra e a Lei de Stefan-Boltzmann (Marshall e Schott, 1999).

A energia irradiada pelo Sol percorre distâncias diferentes na atmosfera até ser absorvida nesta ou na superfície da Terra. Por esta razão e devido à sua forma, movimento de rotação e translação, e eixo de inclinação da Terra, a radiação solar incide na superfície da Terra de forma assimétrica, por efeito de vários fatores como: estação do ano, latitude, hora do dia, estado do mar <sup>1</sup> e nebulosidade <sup>2</sup>, que influenciam o ângulo de incidência da radiação solar e consequentemente, a sua intensidade no local (Sælen e Aas, 2012; Stewart, 2008).

A energia transferida entre um sistema com temperatura mais alta para outro de temperatura mais baixa denomina-se calor. A temperatura é a medida do nível de agitação médio entre moléculas e macroscopicamente, é a propriedade que determina se um sistema está em equilíbrio térmico <sup>3</sup> com outros (Stewart, 2008).

---

<sup>1</sup>Altera a refletividade da superfície.

<sup>2</sup>A sua ausência causa transparência da atmosfera.

<sup>3</sup>Dois sistemas estão em equilíbrio térmico quando estão à mesma temperatura.

O excesso de calor incidente na superfície do mar nos meses de maior intensidade de radiação, como no verão, é compensado pela ausência deste nos meses de inverno. A energia é armazenada no oceano, nas camadas até aproximadamente 100 metros e retorna à atmosfera nos meses de inverno. Devido a esta capacidade de armazenamento de energia pelo oceano, a TSM sofre alterações menores do que a superfície terrestre (Gill, 2016; Stewart, 2008).

A diferença da intensidade da radiação solar nas latitudes baixas (onde é mais elevada) e nas latitudes elevadas (onde é mais baixa) provocam instabilidade térmica. Por forma a atingir-se o equilíbrio térmico, existem mecanismos naturais de compensação que transferem o calor das baixas para as elevadas latitudes: os ventos e as correntes. Assim, é induzido movimento no sistema atmosfera-oceano. Estes processos de interação deste sistema influenciam o estado do tempo e o clima, provocando alterações na TSM (Gill, 2016; Stewart, 2008).

Estudos anteriores revelam que não se consegue descrever variações nas correntes ou prever efeitos a longo prazo das mudanças climáticas à superfície até se estabelecer as ligações entre a superfície do oceano nas diferentes estações do ano. Também demonstram que o aquecimento global crescente no século XXI tem causado um aumento de TSM, o que altera as condições oceanográficas (Gibelin e Déqué, 2003; Giorgi, 2006; Zhou et al., 2000).

A TSM não é um parâmetro trivial de descrever, nomeadamente pela dificuldade em delinear os limites que definem a superfície. Sensores com sensibilidade a diferentes comprimentos de ondas eletromagnéticas medem profundidades diferentes na camada de água (Brennan, 2004; Emery et al., 2001; GHRSST, s.d.). Assim, a profundidade da TSM irá variar conforme o método de medição utilizado:

- Sensores de radiação infravermelhos são sensíveis à radiação com um comprimento de onda entre 3,7 e 12  $\mu\text{m}$  e medem temperaturas numa profundidade entre 10-20  $\mu\text{m}$  na coluna de água. Esta profundidade foi denominada *skin*;
- Sensores de radiação micro-ondas detetam radiação com um comprimento de onda entre 2.7 e 5 cm. A esta profundidade é atribuído o nome de *sub-skin*;
- Sensores de contacto *in-situ* apenas conseguem medir a camada imediatamente inferior à *sub-skin*: a camada é denominada *depth*.

Em suma, os sensores infravermelhos medem a TSM na camada denominada *skin*; os sensores micro-ondas medem a TSM da camada imediatamente inferior,

chamada *sub-skin* e por último, as boias derivantes medem a temperatura na camada a seguir: a  $TSM_{depth}$ .

A temperatura da camada de superfície, entre a fronteira ar-mar e os 10 metros de profundidade, apresenta uma variabilidade diurna e sazonal em resposta à estratificação térmica mais ou menos acentuada, de acordo da turbulência existente no local (Stewart, 2008). Na figura 1, Donlon et al. (2009) contrapõe dois perfis de temperatura: (a) condições diurnas e noturnas com ventos fortes e (b) condições diurnas com ventos fracos e elevada intensidade de radiação solar. Esta figura permite verificar que:

- O aumento de turbulência provoca uma maior mistura de água e consequentemente diminui a estratificação térmica e a variação térmica ao longo da camada de superfície;
- Esta variabilidade altera-se ao longo dos 10 metros (entre  $TSM_{int}$  e  $TSM_{fnd}$ ). Nos dois gráficos, as sub-camadas *int*, *skin*, *sub-skin* têm maior variação térmica do que a sub-camada *depth*. A sub-camada *fnd* não sofre nenhuma alteração.

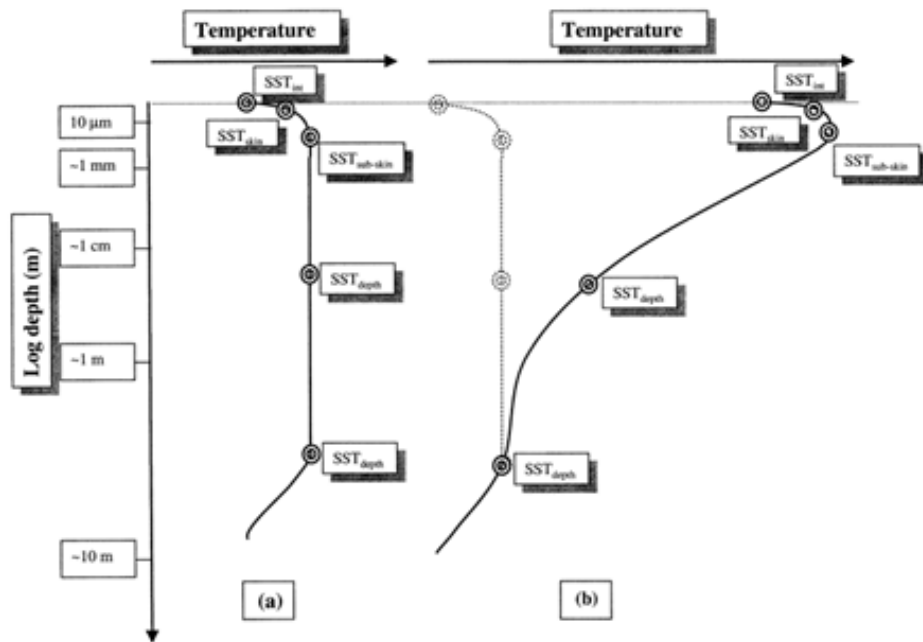


FIGURA 1: Perfis de temperatura da camada próxima à superfície (aproximadamente até 10 m de profundidade) do oceano - (a) condições diurnas e noturnas com ventos fortes; (b) condições diurnas com ventos fracos e elevada intensidade de radiação solar (retirada de Donlon et al., 2001).

A medição da TSM é possível através de diferentes metodologias associadas a diferentes plataformas. As diversas metodologias de medição da TSM apresentam

diferentes características, como, por exemplo, a profundidade da camada da superfície, e apresentam limitações. Como tal, cada metodologia cobre determinadas dimensões espaciotemporais e, portanto, é preciso reconhecer quais são as condições de cada metodologia e qual se adequa melhor para resolver determinado processo (Emery et al., 2001; Matthews, 2013):

- Navios, o método mais ancestral. Esta plataforma pode estar associada a dois tipos de medições: por balde ou por injeção nos motores. O primeiro consiste em recolher água na superfície do mar e com um termómetro de mercúrio e medir a temperatura da água (pode-se recolher água até dezenas de centímetros). Já o segundo, traduz-se na medição da temperatura da água que sofre sucção e entra no circuito de água salgada de refrigeração que transporta a água do mar para os motores diesel propulsores, para arrefecimento dos mesmos. Hoje em dia, os termístores são inseridos na quilha do navio por forma a diminuir o viés positivo que se verifica nos dados derivados do aumento da temperatura associado ao calor provocado pelos motores;
- Plataforma fixa (boias fundeadas) é outro método tradicional. Este efetua registos pontuais, mas com séries temporais longas tendo por isso aplicações operacionais (assimilação de modelos) e climáticas. Emery et al. (2001) demonstrou que a utilização de boias derivantes e fundeadas pode ser combinada, embora as fundeadas meçam não só na superfície como nas camadas mais profundas;
- Plataformas derivantes (Boias derivantes), em que cada boia efetua também registos pontuais enquanto alteram as suas posições. Deste modo, têm pouca uniformização pois todas as boias se encontram em movimento. Programas mais antigos podem ter series temporais longas e cobrir uma área espacial grande. Por esta razão, um dos objetivos do *Global Drifter Program* (GDP) é manter a matriz de boias por forma a garantir a variabilidade espacial e temporal de maneira a conseguir ter também aplicações operacionais (como, por exemplo, assimilação de modelos) e climáticas. Uma vez lançadas, não é possível calibrar os sensores de temperatura.
- Satélite, este efetua registos globais com frequência temporal grande relativamente aos anteriores. Possui também diversas aplicações operacionais (como assimilação de modelos) e estudo de processos.

Guan e Kawamura (2004), Donlon et al. (2007) e Gentemann et al. (2009) enfatizam a dificuldade de comparar o TSM de diferentes plataformas, dadas as

suas características específicas. Nestes estudos é demonstrada a dificuldade em generalizar os resultados de cada um desses trabalhos para determinada zona do oceano e para uma determinada aplicação, pois está muito dependente de fenómenos de mesoescala e condições meteo-oceanográficas locais.

Adicionalmente, em 2011, foi efetuado um estudo comparativo entre a temperatura do mar *in-situ* e satélite no mar Mediterrâneo ocidental para o ano de 1999 em que foi demonstrado que os resultados da avaliação do erro variam com o tipo de sensor, a profundidade das medições *in-situ* e a base de dados utilizada (Alvera-Azcárate et al., 2011).

O objetivo principal do presente estudo é analisar e comparar três fontes de dados atuais de temperatura à superfície, discutindo as suas vantagens e limitações para diversas aplicações, nomeadamente aplicações navais. No seu seguimento, foram identificados outros quatro objetivos que irão servir de suporte para o objetivo principal:

1. Aprofundar os conhecimentos sobre processos que influenciam a TSM e as principais correntes de superfície e suas características;
2. Analisar os três métodos de obtenção de dados TSM por forma a investigar as suas particularidades;
3. Constatar as eventuais diferenças entre dados de temperatura de superfície do mar obtidos pelas boias derivantes do GDP ( $TSM_{GDP}$ ), dados de temperatura de superfície do mar obtidos pelo modelo numérico em estudo ( $TSM_{MOD}$ ) e dados de temperatura de superfície do mar obtidos pelas medições efetuadas por satélites ( $TSM_{SAT}$ ), e discutí-las;
4. Identificar vantagens e limitações no uso dos três métodos em estudo.

A motivação deste trabalho surge no seguimento da colaboração da Marinha Portuguesa e do Instituto Hidrográfico (IH) no GDP da *National Oceanographic and Atmospheric Administration* (NOAA). Os navios portugueses em associação com o IH, todos os anos lançam boias derivantes. Em especial, no ano 2020, o Navio da República Portuguesa (NRP) Sagres lançou diversas boias no decorrer da sua missão de volta ao mundo.

A dissertação é composta por três capítulos. O primeiro capítulo – Dados e métodos – é constituído pelos subcapítulos de descrição dos dados, métodos estatísticos aplicados e tratamento dos dados. O segundo capítulo – Análise de resultados – tem integrado os subcapítulos de análise da TSM na Macaronésia através dos dados

GDP, análise comparativa da TSM proveniente das três fontes e a análise das boias derivantes lançadas pelo NRP Sagres. No terceiro capítulo é efetuada a discussão dos resultados.

## Caracterização da região da Macaronésia

Foi delimitada uma região por forma a incluir todas as ilhas da Macaronésia, isto é, ilhas dos Açores, Madeira, Canárias e Cabo Verde. Através das coordenadas geográficas dos limites desta região, que se encontram representados na tabela 0.1, estabeleceu-se a área de investigação compreendida entre as latitudes 14° e 40°N e a longitude entre 8° e 32°W (Figura 2).

	<b>Latitude</b>	<b>Longitude</b>
<b>Limite superior e ocidental no Arquipélago dos Açores (Ilha das Flores)</b>	39° 27.1' N	031° 16.1' W
<b>Limite oriental nas Ilhas Canárias (Lanzarote)</b>	29° 12.2' N	013° 25.13' W
<b>Limite inferior nas Ilhas de Cabo Verde (Ilha Brava)</b>	14° 48.3' N	024° 42.1' W

TABELA 0.1: Limites geográficos da área de interesse.

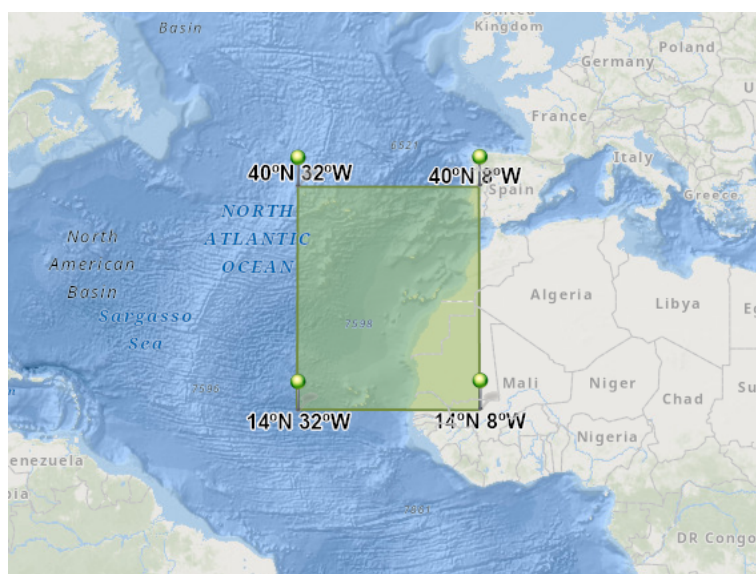


FIGURA 2: Mapa geográfico da área de investigação.



Na região em estudo, o fenómeno de larga escala responsável pela circulação de superfície é o giro subtropical anticiclónico do Atlântico Norte cuja influência é observável a sul dos 45°N (Mann, 1967). Este é forçado pela tensão do vento que ocorre horizontalmente <sup>4</sup> nas primeiras centenas de metros de profundidade (camada de Ekman) e pelo efeito de *Coriolis* que faz desviar 30° para a direita do vento no *Hemisfério Norte* (HN) (Stewart, 2008). Na área de estudo, estamos perante os ventos alísios dos 14° aos 30° N e perante os ventos de Oeste dos 30° aos 40°N .

Um sinal da circulação de superfície sazonal da zona tropical nordeste do Atlântico em larga escala é resposta oceânica à variabilidade sazonal dos ventos e ao deslocamento meridional na *Zona de Convergência Intertropical* (ZCIT). Esta zona está fortemente relacionada com a variabilidade do *upwelling* costeiro forçado pelo movimento Norte-Sul do vento, conforme a estação do ano (Fernandes et al., 2005).

A interação entre o *upwelling* costeiro e os movimentos oceânicos que ocorrem *offshore* resulta em fronteiras irregulares (zona de transição) entre as águas frias e ricas em nutrientes provenientes das camadas mais profundas e as águas costeiras mais quentes e pobres em nutrientes (Barton et al., 1998).

O atrito entre o vento e a água transfere energia provocando o movimento desta. Caso este vento seja persistente, desenvolve-se a corrente entre 20° a 60° para a direita da direção do vento no HN. A batimetria provoca o redireccionamento das correntes na proximidade dos continentes pelo que as isotérmicas se desviam para Norte e para Sul. As correntes equatoriais são redireccionadas para os polos ao longo da costa até se unirem com as correntes para Oeste nas latitudes médias (Gill, 2016).

As maiores correntes de grande escala associadas à região leste do giro do Atlântico Norte são a *North Atlantic Current* (NAC), a *Azores Current* (AzC) e a *Canary Current* (CaC) (Figura 4). Estas correntes têm cerca de 35, 12 e 5 Sverdrup <sup>5</sup> (Stramma 2001). A NAC é uma corrente de escala da bacia localizada no norte da Península Ibérica entre 48° e 53°N (Mason et al., 2005).

A circulação atmosférica da alta pressão dos Açores traduz-se em ventos de Norte/Noroeste ao longo da costa da Península Ibérica, induzindo o fenómeno de *upwelling*, normalmente começando em maio e terminando no mês de setembro (Mann, 1967). Associado a este, surge a Nortada que é um regime de ventos de

---

<sup>4</sup>Os ventos causam movimentos horizontais, mas podem também resultar em movimentos verticais.

<sup>5</sup>Unidade de medida de transporte das correntes oceânicas em volume. Equivale a 106 metros cúbicos por segundo.

mesoescala, presente na costa oeste da Península Ibérica durante todo o ano, principalmente nos meses de Junho, Julho e Agosto. É causada por um regime de alta pressão presente na região dos Açores e por uma baixa pressão continental de origem térmica, que gera circulação de ventos de Norte, paralelos à costa. Assim, durante o Inverno, devido à alta dos Açores estar localizada mais a sul (cerca de 30°N), existe um gradiente significativo das temperaturas na direção norte-sul que se traduz em ventos de velocidade e direção variável enquanto no Verão, pelo facto da alta dos Açores se encontrar mais a norte (cerca de 35-40°N), as direções de vento são de norte/noroeste, em que há grande intensidade do mesmo, devido ao contraste de temperatura entre a massa continental e o oceano (Chaves, 2015).

Entre o sul da costa de Portugal e a costa norte de Marrocos (33°-37°N), o *upwelling* é intermitente e de curta duração. A sul de 20°N, o *upwelling* ocorre maioritariamente nos meses de inverno (Barton, 2001).

O noroeste de África é caracterizado como uma região de *upwelling* que difere das restantes devido à presença das sete ilhas principais do Arquipélago das Canárias, que fornecem outra fonte importante de atividade de mesoescala. Estas encontram-se na zona de transição entre a região de *upwelling* da plataforma continental do noroeste da África e as águas de oceano aberto do giro subtropical. Deste modo, são uma barreira tanto para o fluxo da CaC em direção ao equador e ao fluxo gerado pelos ventos alísios, promovendo a ocorrência de fenómenos de mesoescala. Os fenómenos de mesoescala são aqueles cuja escala espacial se encontra entre aproximadamente uma e dez vezes o raio de Rossby, com uma escala temporal concomitante de um a dez períodos de inércia. Esta variação está associada a *eddies* ciclónicos ou anticiclónicos (Barton et al., 1998).

Cabo Verde encontra-se na fronteira leste do giro, no limite sul da CaC. Esta corrente transporta água fria para Sul ao longo da costa africana e desvia para oeste da costa entre as latitudes 20° e 25°N e gradualmente torna-se na *North Equatorial Current* (NEC) (Fernandes et al., 2005).

A principal característica da região de Cabo Verde é a Zona Frontal de Cabo Verde, separando águas de origem tropical a Sul e de origem subtropical a Norte. Dessa forma, interceta a encosta a norte do Cabo Branco entre 22° e 23°N (Peña-Izquierdo et al., 2012).

A AzC flui para sul à medida que se aproxima da fronteira oriental e onde se formam dois ramos separados pelo arquipélago da Madeira. Varia sazonalmente sendo que oscila para sul no verão e para norte no inverno de forma coordenada

com a oscilação da CaC. Ao longo da AzC, verificam-se vários fenómenos de *eddies*, sendo a sua formação devida à instabilidade baroclínica (Barton, 2001; Brügge, 1995; Mason et al., 2005; Teles-Machado et al., 2007; Zhou et al., 2000). A contracorrente dos Açores localiza-se entre 35° e 36°N, a norte da Corrente dos Açores, iniciando na fronteira leste e estendendo meridionalmente por 110 km<sup>6</sup>. Esta é uma corrente que se desloca para oeste impulsionada pelos *eddies* criados pela AzC e pela beta-pluma topográfica e recebe uma contribuição da Corrente de Portugal. Nas camadas intermédias, esta corrente transporta água proveniente do mar Mediterrâneo na extremidade oeste do Estreito de Gibraltar, normalmente com uma intensidade de 1 Sv, aumentando conforme desce a encosta continental (Brügge, 1995; Volkov e Fu, 2010). O seu transporte é sazonalmente constante, embora haja variabilidade da circulação geostrófica: no verão é mais forte perto da costa africana enquanto no inverno é mais forte no oeste das ilhas Canárias (Mason et al., 2005).

Por sua vez, a corrente Norte Atlântica segue para sul originando a *Portugal Current* (PC). Esta corrente, lenta e superficial, flui para sul ao largo da Península Ibérica, e separa-se na costa africana a cerca de 20°N tornando-se a corrente Norte Equatorial (NEC). A sul deste ponto de separação existe uma célula de recirculação em torno da Guiné, a leste das ilhas de Cabo Verde (Barton, 2001; Jia, 2000; Mason et al., 2005; Volkov e Fu, 2010) (Figura 3). Esta possui uma contracorrente em profundidade no sentido dos pólos, a norte de 41°N, durante o outono, inverno e início da primavera e é a corrente de ligação entre NAC e AzC e tem um transporte de, aproximadamente, 3 Sv (Relvas et al., 2007).

A corrente das Canárias, que flui pelas ilhas Canárias (Pérez-Hernández et al., 2013), é caracterizada por uma estrutura de mesoscala intensa na zona de transição entre as águas frias e ricas em iões provenientes do *upwelling* costeiro e as águas mais quentes e oligotróficas do oceano aberto. Esta corrente causa fronteiras irregulares e forma fluxos em alto mar de água de *upwelling*. (Barton et al., 1998). Os efeitos combinados da perturbação do fluxo, *upwelling* de fronteira entre as águas de *upwelling* e oceânicas produzem um padrão complexo de variabilidade regional, com a presença de *eddies* ciclónicos e anticiclónicos a junto ao Arquipélago das Canárias (Figura 4). Esta corrente varia sazonalmente em ligeiras mudanças de posição, ficando mais próximo da costa africana no verão e mais longe no inverno.

A corrente Norte Equatorial (NEC) cruza a costa africana e segue para sul num fluxo estreito e intenso em direção ao equador, situando-se maioritariamente na

---

<sup>6</sup>Aproximadamente 59,4 milhas náuticas.

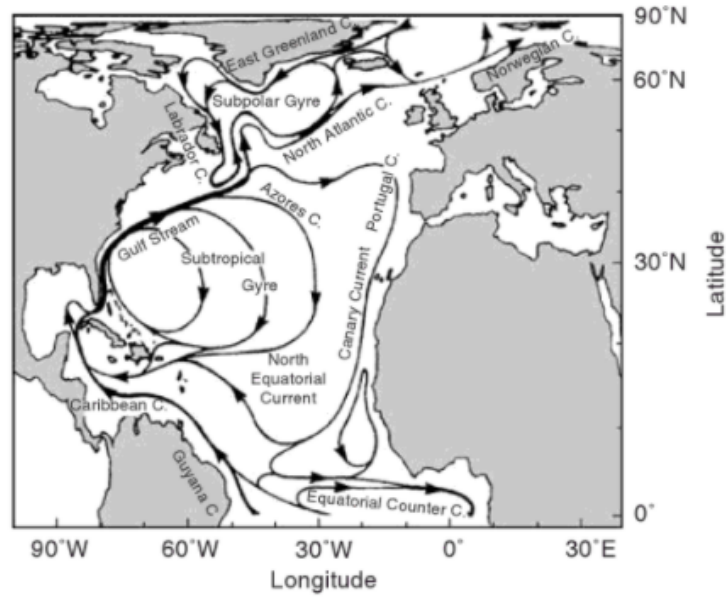


FIGURA 3: Esboço da circulação geral próxima à superfície do Oceano Atlântico Norte (retirado de Barton, 2001).

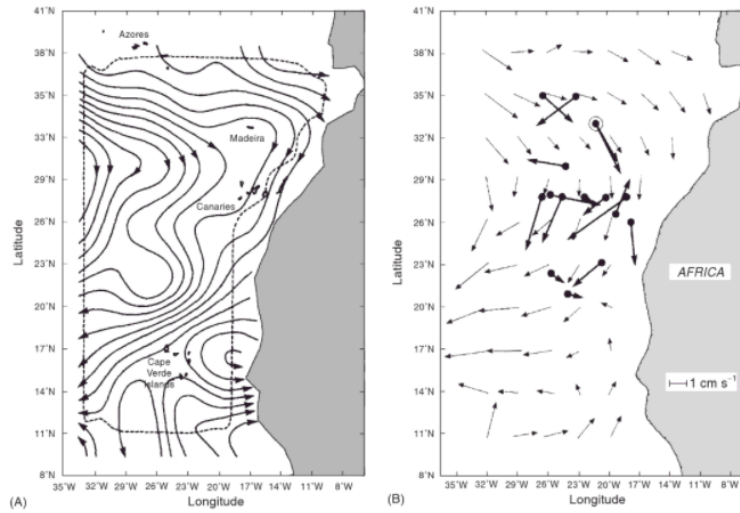


FIGURA 4: Volume de transporte total do campo médio de densidade de longo prazo com assunção geostrófica e até 200m de profundidade(A) e vetores de corrente geostrófica calculado a partir do campo de densidade média de longo prazo e correntes médias observadas (B) esperados para a região em estudo (retirado de Barton, 2001).

faixa costeira entre 20°N e 30°N, particularmente na primavera e verão. A recirculação em torno da Guiné é mais clara no inverno e na primavera, mas as indicações de fluxo para Norte também são verificadas noutras estações. No outono, oscila para Norte estendendo-se até as Canárias.

Para a comunidade científica, a qualidade das medições das propriedades da água do mar é de grande importância por permitir validar modelos. A precisão<sup>7</sup> e exatidão<sup>8</sup> dos dados têm sido proporcionais ao avanço tecnológico. Comparar dados de diferentes fontes é importante para a intercomparabilidade de resultados obtidos por diferentes equipas de investigação e para a compreensão das tendências de variabilidade temporal e espacial (Brügge, 1995; Lauvset e Tanhua, 2015).

---

<sup>7</sup>Diferença entre contínuas medições (Stewart, 2008).

<sup>8</sup>Diferença entre o valor medido e o valor real (Stewart, 2008).



# Capítulo 1

## Dados e Métodos

Neste capítulo são apresentados, em três seções distintas, a descrição detalhada dos métodos que originaram as séries de dados utilizadas no estudo, os métodos estatísticos aplicados na análise e comparação de cada conjunto de dados e, por último, o processo de tratamento dos dados

### 1.1 Descrição dos dados

Os dados em estudo foram retirados de três metodologias diferentes: boias derivantes do GDP, um modelo numérico de previsão e análise, e um produto L4 de satélites, ambos no âmbito do *Copernicus Marine Environment Monitoring Service* (CMEMS). Os três conjuntos de dados foram extraídos de forma *online* e gratuita, para a região e período em estudo.

#### 1.1.1 Dados das boias derivantes do GDP

O GDP é um programa científico criado pela NOAA em 1979, com o intuito de estudar o clima mundial, através de uma matriz global de 5° por 5° de boias derivantes, aproximadamente 1300 boias, (constituída atualmente por 1419 boias) lançadas em todo o mundo (IH, s.d.; NOAA, s.d.-c).

O seu objetivo principal é melhorar a previsão numérica global e a monitorização dos oceanos e do clima, através de boias derivantes. As boias adquirem dados de posicionamento e temperatura, podendo também medir pressão atmosférica, ventos e ondas. Os dados adquiridos pelas boias são transmitidos, em tempo real, por satélite, processados e inseridos no Sistema de Telecomunicação Global (GTS), para distribuição e uso mundial (Figura 1.1). Os dados são usados para a previsão numérica, monitorização dos oceanos e do clima, investigação científica e

segurança da navegação (IH, s.d.; NOAA, s.d.-c).

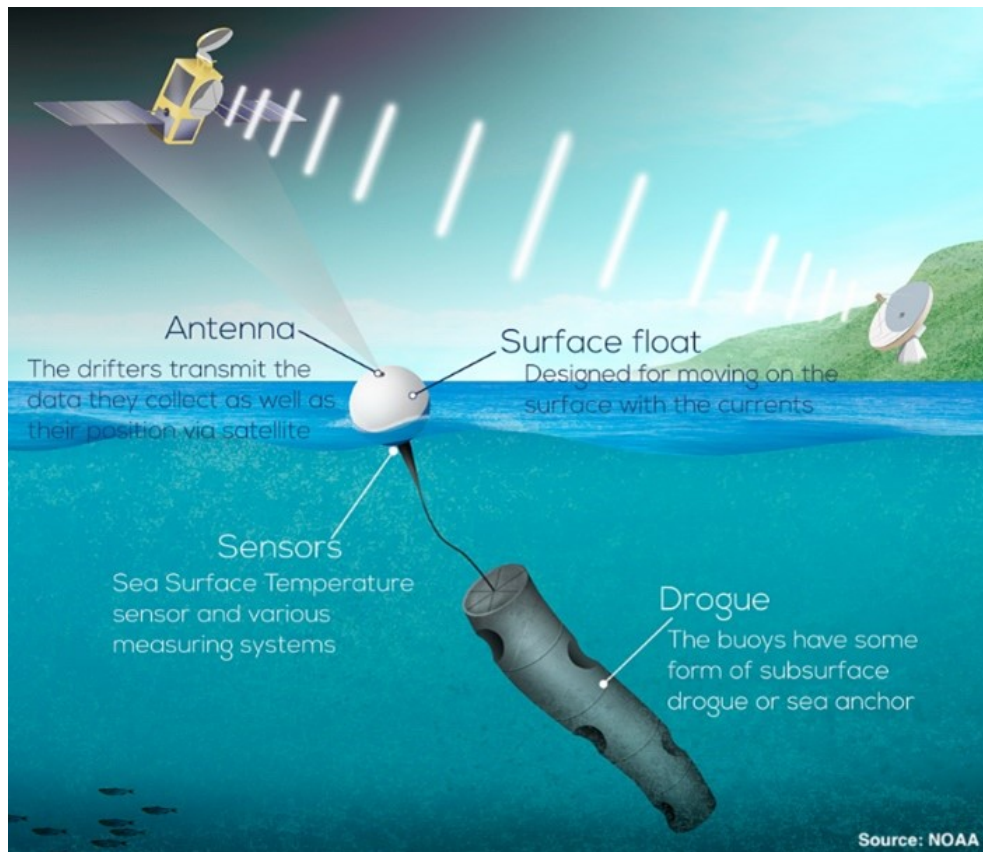


FIGURA 1.1: Esboço do funcionamento das diferentes partes uma boia derivante (retirada de NOAA, s.d.-c).

Os registros das boias derivantes foram retirados do conjunto de dados “*Global Drifter Program drifting buoy collection*” do servidor ERDDAP disponibilizados online pela NOAA (NOAA, s.d.-a). Assim, foram descarregados os 78316 registros de 114 boias *Surface Velocity Program* (SVP) Iridium e SVP Argos-2 e Argos-3, para um ficheiro .csv.

As boias têm um peso total de aproximadamente 23kg e podem ser divididas em três partes: o flutuador de superfície, o cabo de tensão e o *drogue* (Figura 1.2). Quanto ao flutuador de superfície, este é responsável por os sensores de medição permanecerem à superfície da água e por armazenar as baterias desenhadas para durar pelo menos 400 dias. O flutuador pode ter de diâmetro entre 30,5 a 40 centímetros. Todas as boias deste programa possuem sensores de temperatura (que se encontra a aproximadamente 15 cm de profundidade) e localizador geográfico. Algumas também incorporam barómetro ou outros sensores de salinidade, altura



de onda , velocidade, direção do vento e ainda sensores óticos e de diagnóstico de flutuação de superfície interna (Motyzhev et al., 2012; NOAA, s.d.-b; Yug, s.d.).

As boias SVP têm incorporados dois sensores de temperatura: o sensor analógico que contém o termistor e o sensor de alta resolução, sendo que ambos estão protegidos da radiação solar direta (Le Menn et al., 2019). Os termístores são dispositivos elétricos utilizados para medir temperaturas e são compostos por material semicondutor cuja resistência elétrica varia com a temperatura. Os termístores presentes nestas boias têm sensibilidade a um intervalo de temperaturas entre  $-5^{\circ}$  a  $36^{\circ}\text{C}$  (Argos) e  $-2^{\circ}$  a  $45^{\circ}\text{C}$  (Iridium) uma precisão de  $\pm 0.1^{\circ}\text{C}$  (Motyzhev et al., 2012; Yug, s.d.).

Facultativamente, o flutuador pode estar acoplado a um *drogue* com orifícios abertos, como demonstrado na Figura 1.2. O *drogue* funciona como um peso que confere estabilidade e alguma resistência ao movimento da boia. Este encontra-se centrado nos 15 metros (num intervalo de extensão entre 8 e 20 m) de profundidade e tem um diâmetro de 61 centímetros. Como o *drogue* e o flutuador de superfície estão unidos pelo cabo de tensão, o flutuador também se move com as correntes próximas à superfície. Sem a presença deste, o flutuador será transportado maioritariamente pelo vento e também, em menor dimensão, pela ondulação, continuando a transmitir dados que podem ser descarregados no portal, mas não medem as velocidades das correntes superficiais oceânicas sem serem condicionado pelo vento (Motyzhev et al., 2012; NOAA, s.d.-b; Yug, s.d.).

Neste estudo, apenas foram utilizados os registos que apresentavam *drogues*, por forma a analisar os trajetos que as boias efetuaram, sem os dados estarem totalmente condicionados pelo vento e ondulação.

Cada boia tem um custo base de aproximadamente 1500 euros, mas encarece conforme são adicionados sensores, podendo chegar a mais de 11 mil euros.

As boias têm atualmente uma cobertura geográfica de 77% do globo. Para maximizar a cobertura espacial das boias derivantes, são difundidos mapas de valor de implementação pelos quais os países parceiros se guiam quando planeiam os locais para a lançamento das boias que depende dos seguintes fatores (Machín et al., 2006; NOAA, s.d.-c):

- Número de sensores operacionais nas boias derivantes existentes;
- Presença de *drogue*;
- Tempo de vida da boia derivante a ser implementada;

- Futuros lançamentos de boias derivantes já planeadas;
- Esperança de vida útil das boias derivantes.

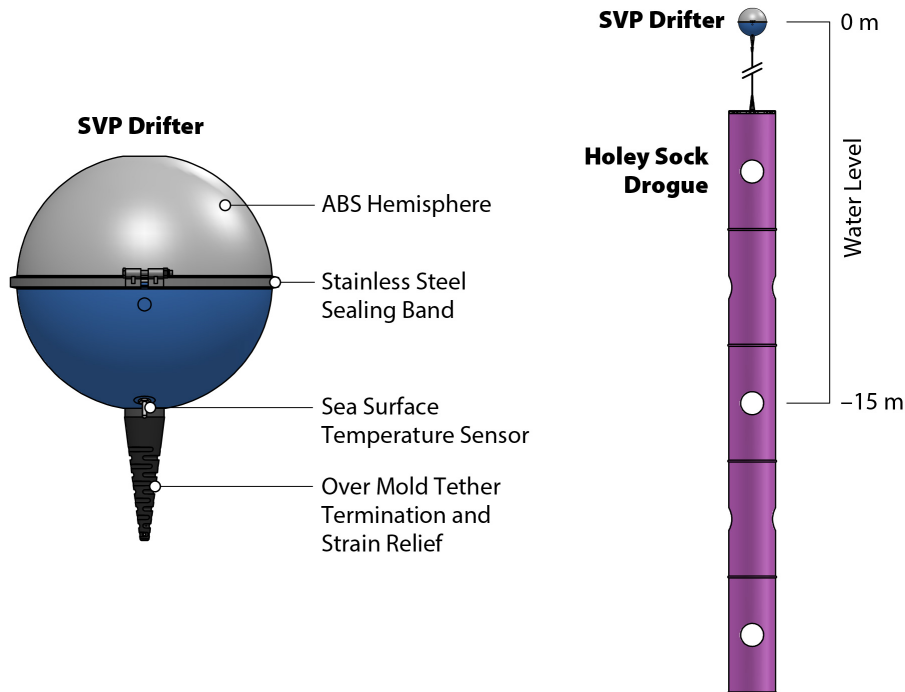


FIGURA 1.2: Esquema de uma boia derivante SVP (retirado do <https://rsaqua.co.uk/product/svp/>).

A área de interesse é uma das áreas de maior valor de implementação no Atlântico Norte (Figuras 1.3 e 1.4). Note-se a elevada importância de lançamento de boias, tanto em janeiro como em junho de 2021, em regiões como entre Portugal continental e os Açores, e na região das Canárias, bem como a oeste destas.

Entre todos os parceiros deste programa, são lançadas anualmente aproximadamente 1000 boias nos oceanos. Portugal é um destes parceiros e vários NRP, nas suas missões têm colaborado no lançamento destas boias que darão dados para o GDP (NOAA, s.d.-c).

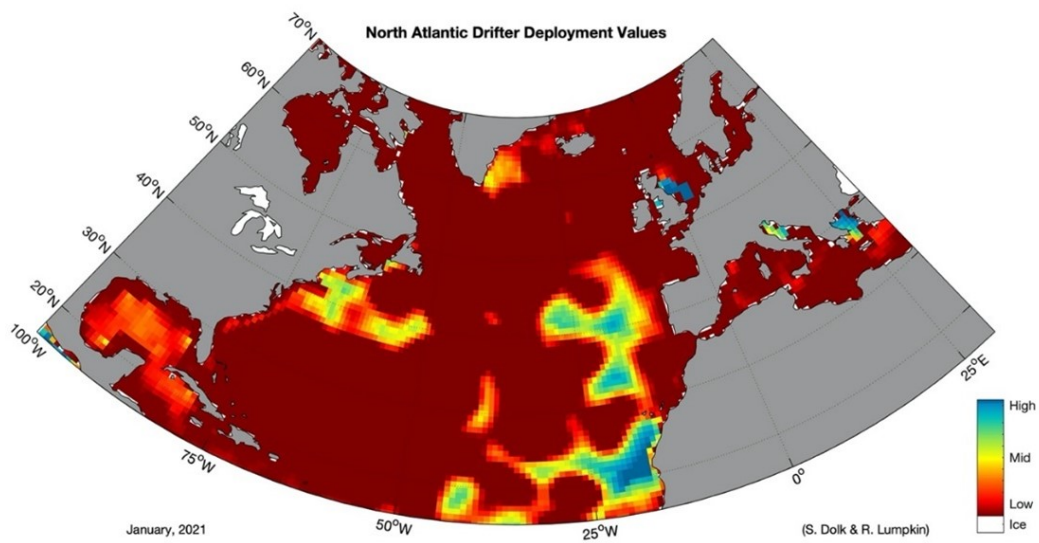


FIGURA 1.3: Mapa de densidade de cobertura de boias de janeiro de 2021 (Retirado de NOAA, s.d.-c).

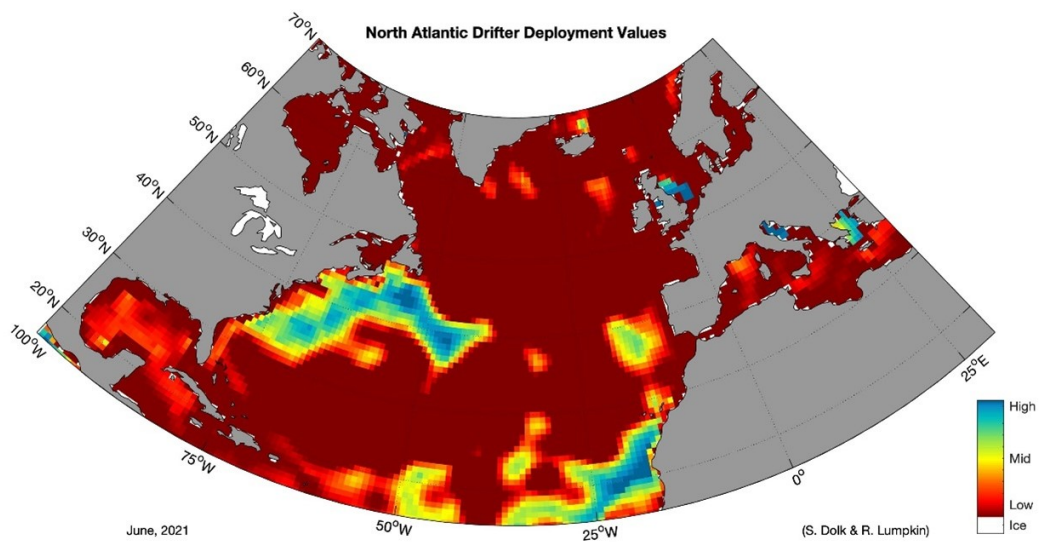


FIGURA 1.4: Mapa de densidade de cobertura de boias de junho de 2021 (Retirado de NOAA, s.d.-c).

### 1.1.2 Dados do modelo numérico

O modelo numérico em estudo é o “Global Ocean  $\frac{1}{4}$  physics analysis and forecast updated daily” (CMEMS, s.d.).

A versão utilizada do modelo foi a V3.1 que foi validada em setembro de 2020. Este é um produto de análise e previsão para vários níveis verticais, de vários parâmetros oceanográficos. Os dados TSM que este modelo fornece são de campos horários instantâneos em formato Network Common Data Form 4 (NetCDF-4). O modelo utiliza o método de assimilação NEMOVAR de 40 iterações para a TSM utilizando dados de satélites, dados *in-situ* de boias derivantes e dados obtidos por navios (CMEMS, s.d.; Guiavarc’h et al., 2019; Lea et al., 2015).

Este modelo utiliza o sistema global de configuração do oceano (GO5) e na versão 3.4 do sistema de modelagem NEMO com o ORCA025 tripolar com uma malha horizontal com um espaçamento de 28 quilómetros, baseado na configuração desenvolvida pela *Mercator Ocean*. O sistema de coordenadas verticais é baseado em níveis geopotenciais usando o DRAKKAR 75, que fornece uma resolução aumentada próxima à superfície, incluindo camadas de superfície de 1 metro o que permite ter sensibilidade à variabilidade diurna, sem comprometer a resolução em profundidade. Os dados TSM deste modelo são assimilados com (Guiavarc’h et al., 2019):

- Dados de TSM de satélites de nível 2 fornecidos pela *Group for High Resolution Sea Surface Temperature* (GHRSSST) que compreende dados do *Advanced Very High Resolution Radiometer* (AVHRR) (NOAA e MetOp), do AMSR-2 (microondas) e *Visible Infrared Imaging Radiometer Suite* (VIIRS) (infravermelho visível);
- Dados TSM *in-situ* de boias fundeadas, derivantes e registos de navios.

A TSM é mapeada espacialmente através do operador de observação NEMO, que cria o *nearest-time-step* equivalentes ao local de observação utilizando interpolação bilinear na horizontais e cúbica na vertical.

Na validação do modelo, os dados TSM possuíam um viés médio (MBE) de  $-0.018^{\circ}\text{C}$  e uma raiz do erro médio quadrático (RMSE) de  $0.22^{\circ}\text{C}$  (CMEMS, s.d.). Também no estudo de Guiavarc’h et al. (2019), foi encontrado um MBE de  $-0.02^{\circ}\text{C}$ .

### 1.1.3 Dados dos Satélites

Os dados de satélites em estudo são provenientes do “*Global Ocean Sea Surface Temperature Multi Product Ensemble* (GMPE)”, também no âmbito do CMEMS.

Este conjunto de vários produtos GHRSSST (GMPE) foi criado pelo *UK Met Office* e recebe diariamente dados de vários produtos de nível L4 por forma a

produzir um conjunto de dados médios de valores. Os produtos incluídos são : CMC, FNMOC, GAMSSA, MGDST, NAVO K10, OISST.v2 AVHRR, OISST.v2 AVHRR + AMSR, OSTIA, RSS MW / IR, RSS MW, RTG (Martin et al., 2012).

Assim, os dados retirados não tem uma hora específica válida durante o dia, uma vez que tem como contribuições de dados TSM válidos para diferentes períodos de tempo.

Este produto tem um processamento de nível 4 pois como são gerados pela combinação de observações multi-satélite, multi-sensor e dados *in-situ* utilizando métodos otimizados de interpolação, não existe lacunas de dados. Os dados estão em malha regular de  $0.25^\circ \times 0.25^\circ$  de resolução horizontal, com alta resolução global ( $<10\text{km}$ ) e em formato GHRSSST NetCDF (GHRSSST, s.d.).

A necessidade da combinação de vários satélites surge do facto de a radiação infravermelha ser condicionada pela existência de nuvens pois devido ao seu comprimento de onda não consegue penetrar enquanto a radiação micro-ondas já não é afetada pela presença de nuvens. Dados os comprimentos de onda que os satélites operam, são medidas diferentes TSM. Os sensores infravermelhos (3,7 a 12  $\mu\text{m}$ ) irão medir a  $TSM_{\text{skin}}$  enquanto os micro-ondas (2.7 e 5 cm) a  $TSM_{\text{sub-skin}}$  (Brennan, 2004) e *in-situ* mede-se a  $TSM_{\text{depth}}$ . Deste modo, os registos serão uma combinação de todos estes dados.

## 1.2 Métodos estatísticos aplicados

Neste estudo, pretende-se efetuar um sumário estatístico e uma comparação entre os dados provenientes das três diferentes fontes. Para o primeiro, foram calculados os valores mínimos, valores máximos, a sua contagem, a média aritmética, e para avaliar a sua dispersão foi calculado o desvio padrão ( $\sigma$ ). Foi também efetuado o teste da normalidade dos dados de *Kolmogorov-Smirnov*. Para a comparação entre conjuntos de dados, foram calculados o *Mean Bias Error* (MBE), o *Mean Absolute Error* (MAE), o *Root Mean Square Error* (RMSE) dos erros, assumindo os dados provenientes da GDP como os dados reais/observados.

O desvio padrão é uma medida de dispersão de dados em torno da média na mesma dimensão da variável em estudo. Um baixo desvio padrão demonstra que existe uma boa precisão da medição, enquanto um valor alto dessa medida indica o contrário. O seu valor pode ser calculado através de:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (p_i - \mu)^2}{n}} \quad (1.1)$$

O MBE indica a diferença média dos valores previstos e os valores observados. Este destina-se a indicar se existe uma sub-estimação ou uma sobre-estimação (Kleissl, 2013) dos dados observados. Na prática, este erro quantifica o viés do modelo e é dado pela seguinte expressão:

$$MBE = \frac{\sum_{i=1}^n (p_i - o_i)}{n} \quad (1.2)$$

O MAE é uma medida amplamente utilizada para medir o desempenho de modelos de previsão provenientes de estudos de investigação meteorológica e climática. Trata-se da média dos valores absolutos das diferenças entre os valores previstos (pelo produto dos satélites e modelo numérico) e os valores observados (valores *in-situ*)(Chai e Draxler, 2014, p. 1247) e expressa-se pela seguinte equação :

$$MAE = \frac{\sum_{i=1}^n |p_i - o_i|}{n} \quad (1.3)$$

O RMSE é também uma medida padrão de estatística para os mesmos fins que o MAE. Indica a qualidade dos resultados das estatísticas com múltiplas fontes e quantifica o erro total de um estudo. Na prática, este é a soma da variância (associada à precisão) com o viés ao quadrado (associado à exatidão)(Chai e Draxler, 2014) e é dado pela seguinte equação :

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (p_i - o_i)^2}{n}} \quad (1.4)$$

Para cada uma das expressões apresentadas,  $p_i$  será o valor estimado pelo MOD ou SAT,  $\mu$  a média dos valores estimados,  $o_i$  o valor observado pelo GDP e  $n$  o número de observações.

Ambos os erros, MAE e RMSE são calculados para avaliar a precisão de um modelo de previsão, no entanto, o MAE tende a ser menor do que o RMSE pois o último tende a penalizar erros grandes (é mais sensível a *outliers*), porque os eleva ao quadrado.

A correlação de Pearson é uma medida de associação linear entre variáveis. Este varia entre -1 a 1 e quanto mais próximo destes extremos (positivo ou negativo), maior o grau de dependência estatística linear entre as variáveis (valores superiores a 0,7 são associados a uma grande magnitude do coeficiente). Uma das suas premissas é que a distribuição dos valores seja normal apesar de saber-se que a partir do Teorema do Limite Central, em grandes amostras a curva de distribuição é aproximadamente normal, independentemente do formato da distribuição dos dados; o coeficiente de correlação é fortemente afetado pela presença de outliers, e esta correlação obriga a independência das observações (Abdi, 2007, p. 118-119).

O teste de Kolmogorov-Smirnov (K-S) foi aplicado para realizar um teste de normalidade à distribuição dos erros no apêndice B. Através do método dos testes de hipótese, define-se a hipótese nula:  $H_0: X \sim N(\mu, \sigma)$ . Caso se verifique  $H_0$  utilizam-se testes paramétricos, caso contrário utilizam-se testes não paramétricos. A rejeição da hipótese determina-se pelo apuramento do valor crítico de distribuição através do *p-value* da hipótese, isto é, se este for menor do que o intervalo de significância ( $\alpha$ ), então rejeita-se a hipótese nula (Rodrigues, 2011). Através deste, verificou-se a não normalidade na distribuição dos dados, pelo que a utilização da correlação de Pearson tornou-se desadequada por não cumprir com as premissas da normalidade e da independência das observações (Figueiredo et al., 2009).

Ao invés utilizou-se a correlação de Kendall, uma medida de associação entre duas variáveis. O coeficiente, também chamado de  $\tau$  de Kendall avalia o grau de similaridade entre duas variáveis calculando as suas diferenças simétricas. Deste modo, analisa se o valor de uma variável aumenta ou diminui conforme o valor da outra variável. Este calcula as distâncias de diferenças simétricas entre dois conjuntos de pares que representam a combinação binária e após, normaliza a distância de forma a que tome valores entre -1 e 1 (Abdi, 2007).

$$\tau = 1 - \frac{2(d\Delta(P1, P2))}{N(N-1)} \quad (1.5)$$

Para a análise dos dados, utilizou-se o *Spyder*, um ambiente de desenvolvimento integrado (IDE) de linguagem *Python* e o *Excel* (Anaconda, s.d.).

## 1.3 Tratamento dos Dados

Nesta secção, evidenciam-se os procedimentos para manuseamento e tratamento dos dados, e preparação para a sua análise. Estes incluem uma averiguação

da representatividade do conjunto de dados, detecção de lacunas de dados e remoção de anomalias, e descrição das características do conjunto de dados. Como resultado, obteve-se uma única base livre de dados discrepantes e que possui todos os dados registados às 12h UTC (a descrição em detalhe encontra-se no Apêndice A).

Como supramencionado, retiraram-se de três diferentes fontes, dados de TSM registados para o período entre 30 de junho de 2019 e 30 de junho de 2020, na região da Macaronésia. As três bases de dados foram sobrepostas espaciotemporalmente para que pudessem ser comparáveis. Uma vez que os registos GDP são valores instantâneos, medidos *in-situ* a cada 6 horas, os valores SAT resultam de médias diárias registados às 12 horas UTC e os dados MOD de TSM são dados instantâneos e registados ao minuto 40 de cada hora, procedeu-se à homogeneização temporal das três bases de dados, intercetando temporalmente as três bases de dados às 12 horas UTC, assumindo-se que a  $TSM_{MOD}$  ao minuto 11:40 seria correspondente ao valor à hora certa. Simultaneamente, também se procedeu à interceção das posições de cada registo relativo às três bases de dados, sendo que a posição de cada boia, em cada passo temporal, definiu as posições geográficas a considerar para a comparação com o MOD e SAT. Para retirar os valores de latitude e longitude mais próximos com os de referência (GDP), foi utilizado o método do vizinho mais próximo.

A base de dados resultante da interceção mencionada anteriormente, é constituída por 19565 conjuntos de  $TSM_{GDP}$ ,  $TSM_{MOD}$ ,  $TSM_{SAT}$ , latitude, longitude e grupo data-hora. Da amostra total, verificou-se que 0,5% dos registos não apresentavam valores, pelo que foram removidos da base de dados final, resultando em 19468 registos. Faz-se referência a dois casos isolados onde se verificaram erros sistemáticos na boia 4101718 e boia 4201517. No primeiro caso, os dados  $TSM_{GDP}$  apresentavam um deslocamento constante relativamente aos dados  $TSM_{MOD}$ ,  $TSM_{SAT}$ , pelo que se assumiu que o termistor apresentava problemas. Já no segundo caso, os dados registados de  $TSM_{MOD}$  eram semelhantes aos dados de  $TSM_{GDP}$  da boia 4201517, mas os dados  $TSM_{SAT}$  apresentavam, sistematicamente, desvios bastante superiores ao expectável. Posto isto, eliminaram-se todos os registos da boia 4101718, mas mantiveram-se os registos da boia 4201517, pois considerou-se que o problema estaria nos dados  $TSM_{SAT}$  e não nos dados  $TSM_{GDP}$  (mais informação pode ser consultada no Apêndice B). Uma vez removidos os dados da boia 4101718, permaneceram 19119 registos que foram utilizados para a análise comparativa da TSM, descrita no capítulo 2.

O número de boias disponíveis na área de estudo não foi constante ao longo do tempo, conforme se pode verificar na Figura 1.5. Durante o período considerado,



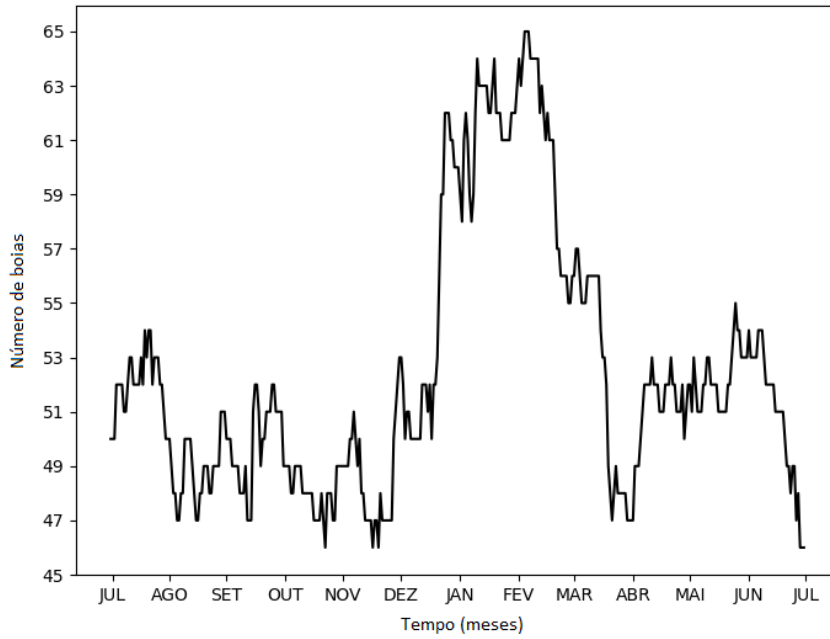


FIGURA 1.5: Variação temporal do número de boias a registrar diariamente, em simultâneo, na área de estudo.

na região da Macaronésia, o número de registos simultâneos de  $TSM_{GDP}$  oscilou entre 46<sup>9</sup> e 65<sup>10</sup>, sendo a mediana 51 boias a registrar em simultâneo. Para todo o período considerado, nunca houve menos de 46 boias a registrar em simultâneo. Houve uma diminuição da quantidade de boias nesta área desde julho 2019 até final do mês de dezembro 2019, quando esta tendência se inverteu havendo um aumento abrupto de 13 boias. No final de março o número de boias a registrar em simultâneo voltou a diminuir, também de forma acentuada. Na Figura 1.6, podemos observar a distribuição do número de boias a registrar diariamente através de um *Boxplot*. Note-se que, 50% do tempo, haviam entre 49 e 54 boias a registrar diariamente na AI. A área de interesse (AI) apresenta uma dimensão de 26° de latitude e 24° de longitude. Em termos espaciais, a maior quantidade de dados foram registados a norte de 27°N e a oeste de 26°W, de acordo com a Figura 1.7.

<sup>9</sup>Ocorreu nos dias 22 de outubro de 2019, 16 e 19 de novembro de 2019 e nos dias 28, 29 e 30 de junho de 2020.

<sup>10</sup>Ocorreu nos dias 4, 5 e 6 de fevereiro de 2020.

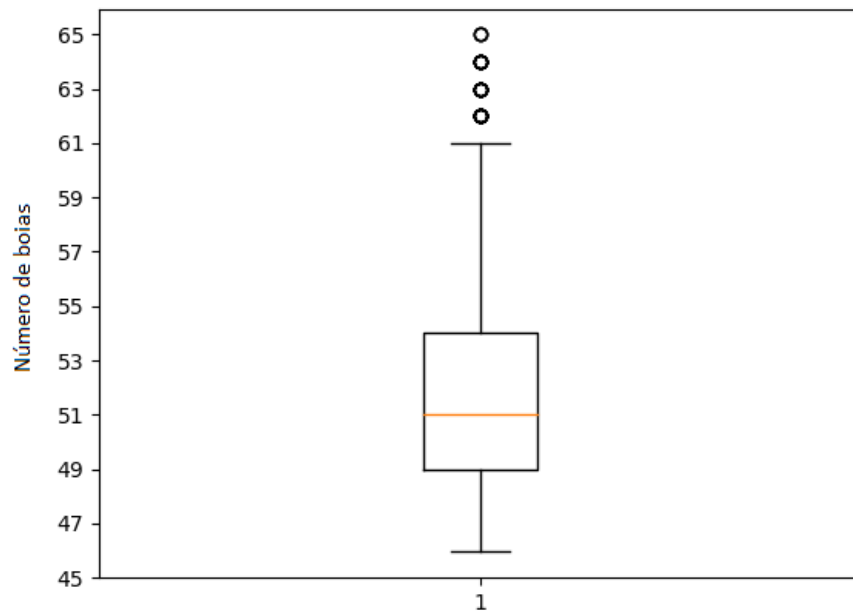


FIGURA 1.6: *Boxplot* representativo do número de boias GDP a registrar diariamente na área de estudo.

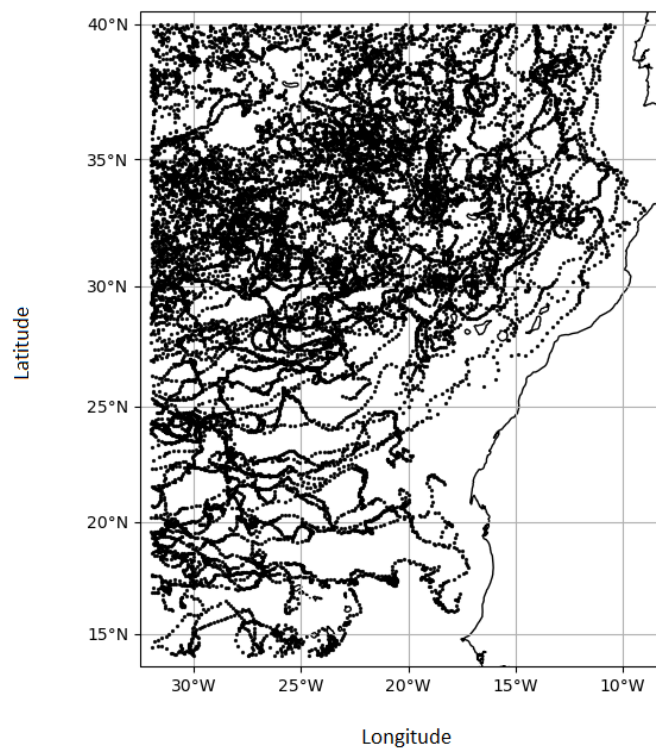


FIGURA 1.7: Posições geográficas das boias GDP na AI, entre 30 de junho de 2019 e 30 de junho de 2020.

Para facilitar a descrição dos dados, a AI foi dividida em 14 subáreas conforme representado na Figura 1.8. Estas subáreas têm dimensão semelhante, que foram escolhidas de maneira que cada subárea fosse representada por uma quantidade de dados GDP mínima. Deste modo, resultaram as subáreas A, B; C, D, L, M e N que apresentam dimensão de 7 graus de latitude por 6 graus de longitude; e as subáreas E, F, G, H, I, J, K com dimensão 6 por 6 graus.

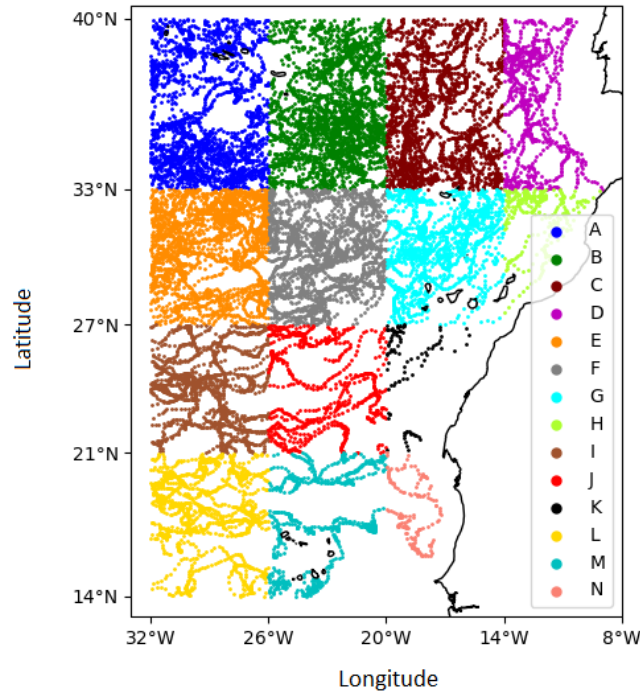


FIGURA 1.8: Posições geográficas das boias GDP nas subáreas, identificadas com as respectivas cores, entre 30 de junho de 2019 e 30 de junho de 2020.

Relativamente à cobertura dos dados, as áreas mais próximas do continente correspondem àquelas com menos registros, isto é, as áreas D, H, K, N. Entre 14 e 21°N e 8 e 14°W não foram registrados dados. Para além disto, existe maior cobertura de dados nas áreas a norte de 27°N e a oeste de 26°W, nomeadamente as áreas B e E. As áreas D, G, H, J, K, L, M e N são as que têm menos registros (possuem dias sem registros). Concluiu-se também que uma área que apresente um maior número de boias não tem necessariamente de ter maior número de registros, de acordo com a Figura 1.9. Por exemplo, a área B tem o maior número de registros, mas o maior número de boias com registros é na área A. Por outro lado, as áreas H e N possuem o mesmo número de boias com registros, mas a área H possui mais 193 registros (Tabela resumo encontra-se no apêndice C).

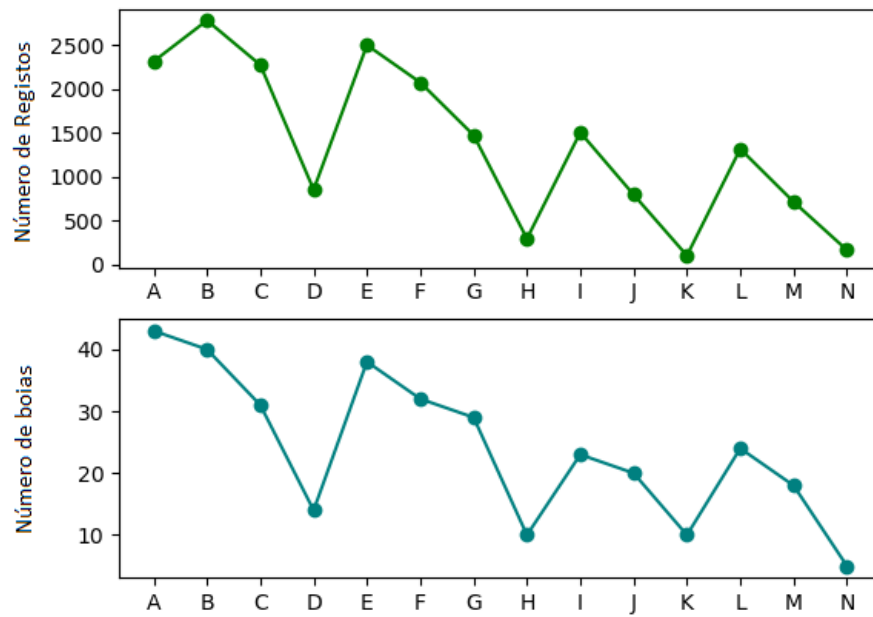


FIGURA 1.9: Quantidade de registos (Em cima) e boias (Em baixo) por subárea.

# Capítulo 2

## Análise dos Resultados

Este capítulo encontra-se dividido em 3 secções, sendo efetuada uma análise aos 19119 registos. A primeira secção diz respeito à análise das TSM provenientes dos dados *in-situ*, a segunda à análise comparativa dos três conjuntos de dados e a terceira à análise das boias lançadas pelo NRP Sagres.

### 2.1 Análise da TSM na Macaronésia através dos dados GDP

Analizando a figura 2.1, onde se apresentam todos os registos diários das  $TSM_{GDP}$ , durante o período em estudo, a TSM foi crescente de julho a setembro e de março a julho, e foi decrescente nos restantes meses.

Analizando a TSM em cada subárea definida no capítulo anterior, verifica-se que aquelas com médias de registos de TSM mais elevadas são a E, I, J, L e N (com médias superiores a 23°C), enquanto as mais baixas são as C, D com uma média anual inferior a 20°C. Quanto à amplitude dos registos, as subáreas A, B, C têm as amplitudes mais elevadas sendo que são superiores a 10°C. Também se comprovou que existem subáreas que não têm registos contínuos durante todo o período de estudo, como é o caso H, J, K, L, M e N (Apêndice C).

Analizando a trajetória das boias que derivaram nos meses de julho 2019, dezembro 2019 e maio 2020 verificou-se a variabilidade sazonal da circulação de superfície que existe nesta região como resposta à variabilidade dos ventos e o deslocamento meridional da ZCIT.

Em julho, ao largo da costa continental portuguesa pode-se observar que todas as trajetórias das boias tiveram direção para sul, por efeito da PC que durante

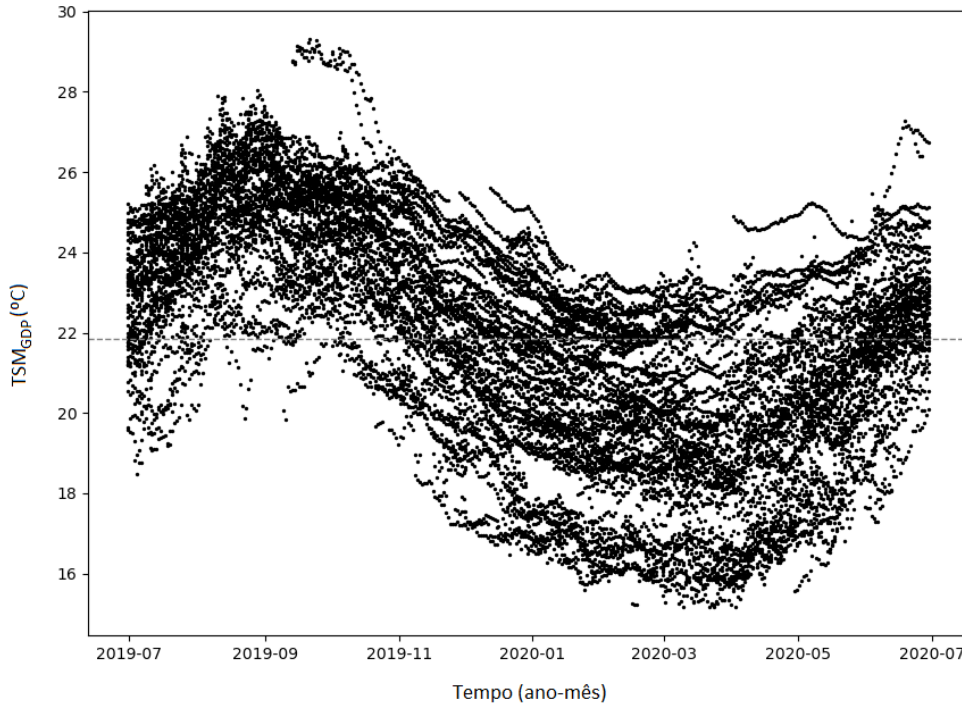


FIGURA 2.1: Variação global temporal dos registos  $TSM_{GDP}$

a primavera e verão tende a encaminhar a circulação para sul e também devido à nortada, que enfatizou esta tendência adicionado movimento para sul/sudeste. Na região mais próxima dos Açores, embora com uma tendência global para a circulação para sudoeste, verificam-se vários movimentos circulatórios nas boias presentes na região, derivados de fenómenos de mesoescala. Ao longo da costa noroeste africana e ilhas Canárias, a CaC encaminhou as boias para sudoeste que, nas latitudes entre 20 a 25°N, sofre um desvio para Oeste, devido a NEC. Na região do Arquipélago da Madeira, as poucas boias existentes no local apresentaram trajetórias para oeste-sudoeste. Por fim, na região a oeste de Cabo Verde, as boias tiveram trajetórias para Oeste (Figura 2.2).

Em dezembro, entre Portugal Continental e os Açores verificou-se que a trajetória das boias era tendencialmente para Oeste, continuando a verificar-se movimentos circulatórios das boias próximo à região do Arquipélago dos Açores. Mais próximo de Portugal Continental existiu uma diminuição significativa da quantidade de boias. Ao longo da costa noroeste africana e ilhas Canárias, não se verificou alteração na advecção das boias, continuando estas a se deslocarem tendencialmente para sudoeste, havendo, no entanto, um desvio para oeste, que se verificou mais a sul, a cerca de 18°N.

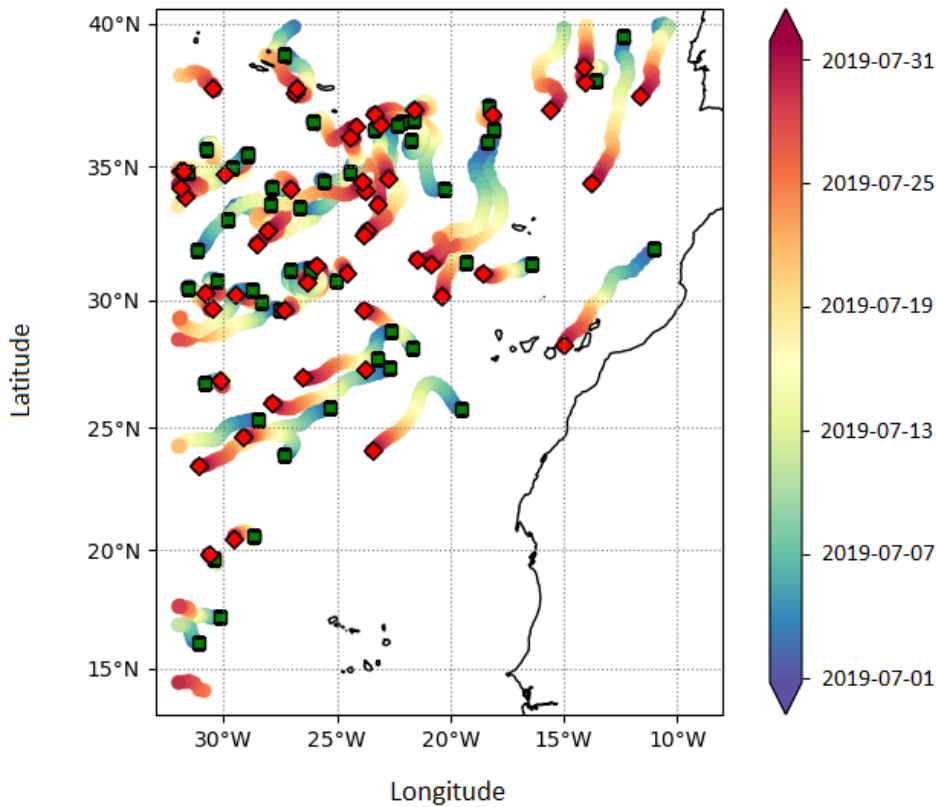


FIGURA 2.2: Posições geográficas das boias GDP entre 1 julho 2019 (losango verde) e 31 de julho 2019 (losango vermelho).

Na região de Cabo Verde, em contraste com o mês de julho, verificaram-se várias boias com movimentos circulatorios, à semelhança do que se verifica junto do arquipélago dos Açores, também no sentido anticiclónico. No arquipélago da Madeira, as boias apresentam um deslocamento consistente para sudoeste (Figura 2.3).

Por último, no mês de maio, a trajetória das boias é no geral semelhante ao mês de dezembro embora apresentem um aumento de movimentos giratórios, possivelmente devido ao aumento de intensidade nos fenómenos de mesoescala relativo à intensidade da circulação média (Figura 2.4).

Utilizando a deriva das boias GDP como uma base para a análise da circulação da região, verificou-se que, como expectável (Barton et al., 1998; Barton, 2001; Brügge, 1995; Fernandes et al., 2005; Jia, 2000; Lauvset e Tanhua, 2015; Mann, 1967; Mason et al., 2005; Peña-Izquierdo et al., 2012; Pérez-Hernández et al., 2013; Relvas et al., 2007; Stewart, 2008; Teles-Machado et al., 2007; Volkov e Fu, 2010; Zhou et al., 2000) a AzC é dominada por fenómenos de mesoescala, tais como *eddies*, estando as boias derivantes presentes nesta região muito suscetíveis

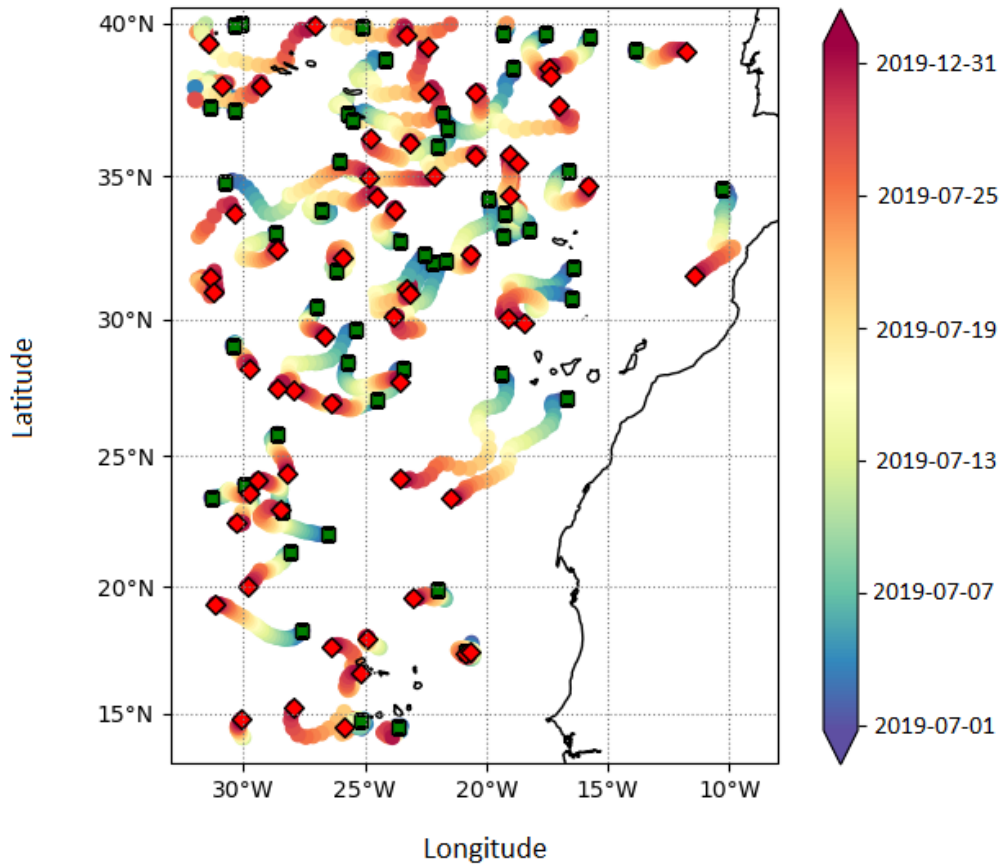


FIGURA 2.3: Posições geográficas das boias GDP entre 1 de dezembro 2019 (losango verde) e 31 de dezembro 2019 (losango vermelho).

a movimentos circulatorios, razão pela qual as boias permaneceram relativamente mais tempo nas subáreas A, B, C, E, F e G. Este facto também se verifica em Cabo Verde (subárea M). Já ao largo de Portugal Continental, Arquipélago da Madeira, costa noroeste Africana e Arquipélago das Canárias, devido às PC e CaC, têm deslocamentos relativamente rápidos, o que resulta em menor tempo de residência nas respetivas subáreas, sendo estas regiões, potencialmente, as mais importantes para lançar boias derivantes. Relativamente à variação sazonal, na corrente de Portugal, verificou-se, nos meses de dezembro e maio, um relaxamento da circulação para sul, resultando na deriva das boias ser dominada pelas estruturas de mesoescala (*eddies*), em vez de derivarem para sul, conforme se verifica no mês de julho. Na região da CaC verificou-se menor variação sazonal, com uma direção consistente ao longo dos meses para sudoeste. A AzC sofreu um relaxamento da circulação para leste nos meses mais frios e verificou-se com maior influência no mês de julho. Na NEC verificou-se uma alteração de latitude, mais a norte nos meses quentes.



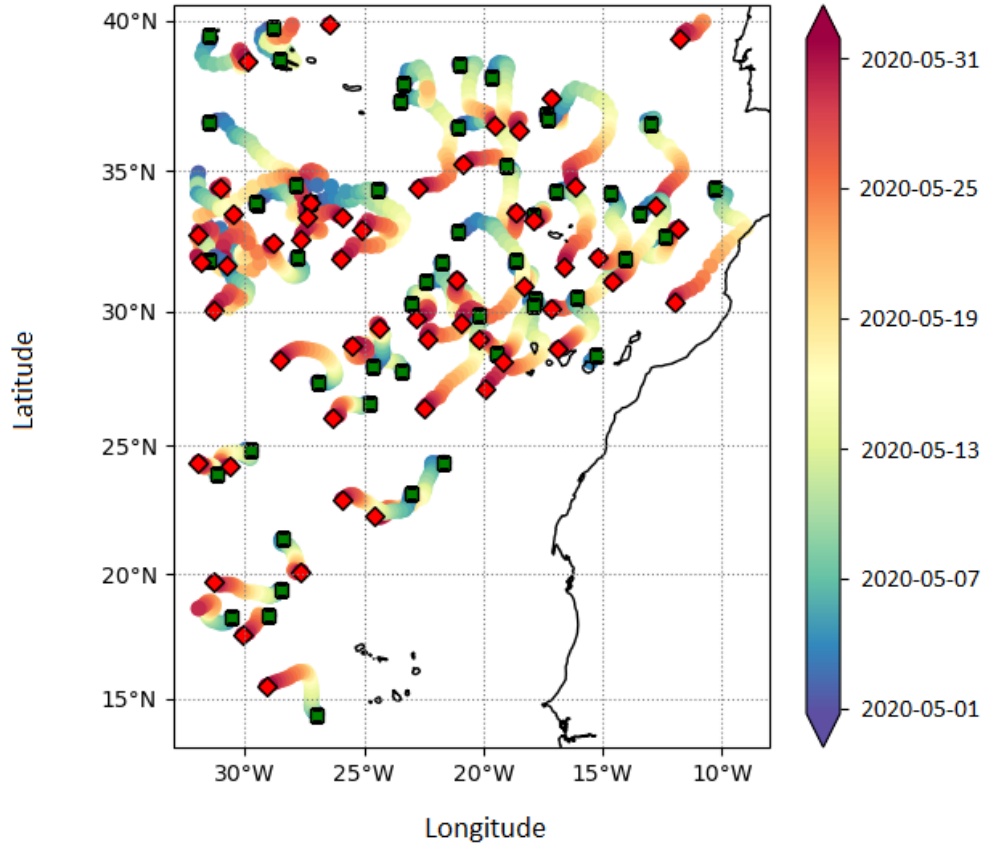


FIGURA 2.4: Posições geográficas das boias GDP entre 1 de maio 2019 (losango verde) e 31 de maio 2019 (losango vermelho).

## 2.2 Análise Comparativa da TSM

Relativamente à diferença entre os valores de TSM obtidos, a grandeza da diferença de valores entre as bases de dados é na ordem das décimas ( $0.1^{\circ}\text{C}$ ). A diferença entre  $TSM_{GDP}$  e  $TSM_{MOD}$  está compreendida entre  $-1.8^{\circ}$  e  $2.6^{\circ}\text{C}$ , numa média de  $-0.1^{\circ}\text{C}$ , e a diferença entre o primeiro e a  $TSM_{SAT}$  foi entre  $-1.8^{\circ}$  e  $4.0^{\circ}\text{C}$ , em média  $-0.1^{\circ}\text{C}$ .

As diferenças entre os valores mínimos, médios e medianas entre  $TSM_{GDP}$  e  $TSM_{MOD}$ , representados na tabela 2.1, são quase idênticos aos valores obtidos entre  $TSM_{GDP}$  e  $TSM_{SAT}$ . Note-se apenas que, é na diferença entre os valores máximos  $TSM_{GDP}$  e  $TSM_{SAT}$  que se regista um maior afastamento.

Observando os valores da matriz de correlação de Kendall obtidos, representada na Figura 2.5, o valor de correlação entre  $TSM_{GDP}$  e  $TSM_{MOD}$  é superior 3 milésimas à correlação entre  $TSM_{GDP}$ , e  $TSM_{SAT}$ .

	GDP e MOD	GDP e SAT
Diferença entre os valores mínimos de cada conjunto de dados	-0,2°C	-0,1°C
Diferença entre os valores máximos de cada conjunto de dados	0,1°C	-0,4°C
Diferença entre os valores médios de cada conjunto de dados	-0,1°C	-0,1°C
Diferença entre os desvios-padrão de cada conjunto de dados	0,0°C	0,0°C
Diferença entre medianas de cada conjunto de dados	-0,1°C	-0,1°C

TABELA 2.1: Resultados da análise estatística da diferença entre os valores de  $TSM_{GDP}$  e cada um dos valores de  $TSM_{MOD}$  e  $TSM_{SAT}$ .

Uma vez que os valores da correlação entre as TSM são bastante semelhantes, apenas separadas por milésima, convém efetuar uma análise mais aprofundada e tentar distinguir cada um dos modelos através dos valores do desvio padrão,  $\sigma$ , do MBE, do MAE e do RMSE.

Quanto ao  $\sigma$ , as  $TSM_{MOD}$  apresentam maior valor, poder-se-ia sugerir que as  $TSM_{SAT}$  apresentavam melhor precisão nas suas medições, no entanto, a distinção é novamente na ordem das milésimas, não sendo de certa forma relevante a diferença registada. O RMSE e o MAE foram superiores na  $TSM_{SAT}$ , pelo que se pode verificar que o seu conjunto de dados é aquele que no total possui mais diferenças entre os valores previstos e os valores observados. Por fim, relativamente ao MBE este foi inferior na  $TSM_{SAT}$  o que indica que, apesar de ambos os métodos terem subestimado os valores de  $TSM_{GDP}$ , a  $TSM_{SAT}$  efetua-o, tenuemente, com uma maior amplitude (Tabela 2.2).

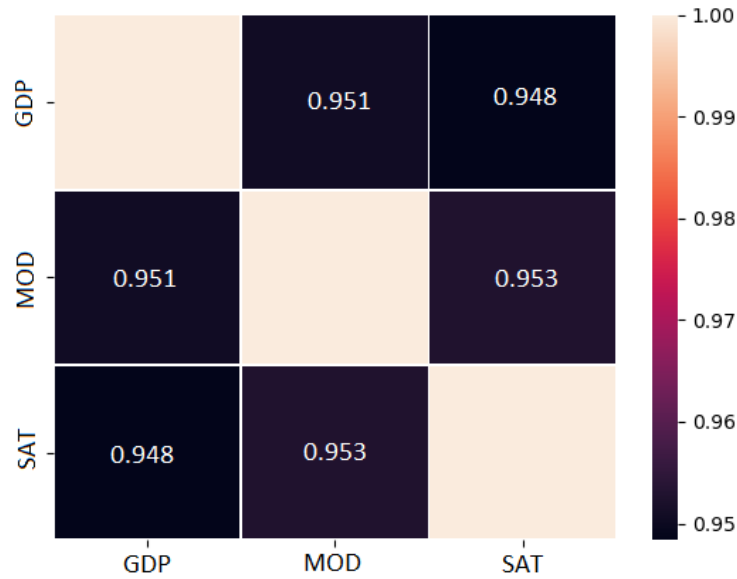


FIGURA 2.5: Matriz de correlação de Kendall entre dados TSM.

$\sigma$	MOD	2,76
	SAT	2,75
RMSE	MOD	0,25
	SAT	0,28
MAE	MOD	0,18
	SAT	0,20
MBE	MOD	-0,06
	SAT	-0,09

TABELA 2.2: Valores de  $\sigma$ , RMSE, MAE e MBE da amostra em estudo.

Numa análise mais aprofundada, poder-se-á comparar a performance dos modelos nas subáreas definidas no capítulo anterior. Comparando os erros verificados nas subáreas e a média global dos erros, apura-se que o MBE do MOD é superior à média global, em valor absoluto, nas áreas A, E, G, K e N enquanto o SAT é nas áreas B, C, D, F, I, J, K, L, N. De realçar que as nas áreas A e E são as únicas em que o MOD subestima mais a realidade do que o SAT (Figura 2.6). Quanto ao MAE do MOD este é superior à média nas subáreas A, B, D, E, K, M, N enquanto o SAT é superior à media nas A, B, E, K, M, N (A subárea D tem um MAE MOD superior ao SAT) (Figura 2.7). Por fim, o RMSE do MOD é superior à média nas

subáreas A, D, E, K, M, N enquanto o SAT é nas subáreas E, K, M e N (As subárea A e D tem um RMSE MOD superior ao SAT) (Figura 2.8). Por outras palavras, as subáreas E, K, M e N tiveram precisões abaixo da média geral nos dois métodos enquanto a subárea A teve abaixo da média geral apenas no MOD.

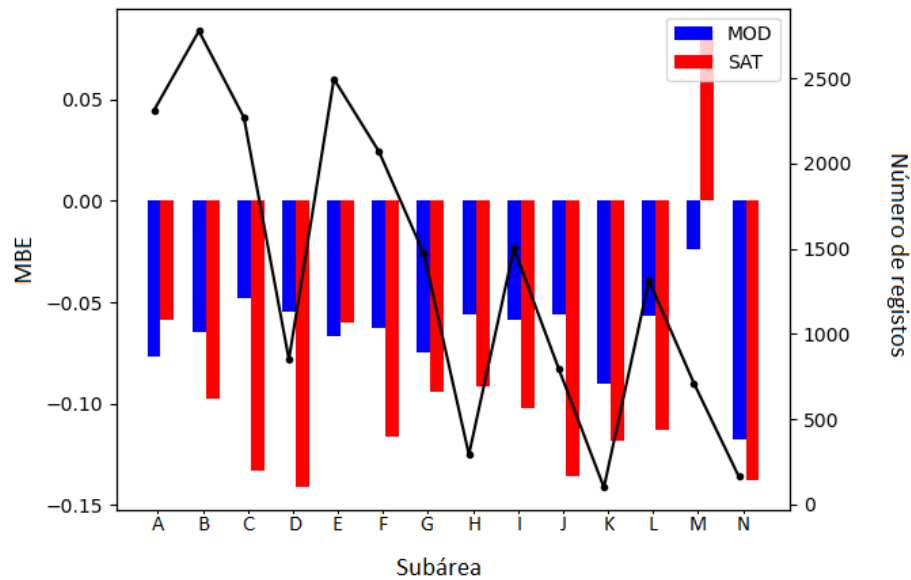


FIGURA 2.6: Valores dos MBE e número de registos por cada subárea definida no capítulo 1.

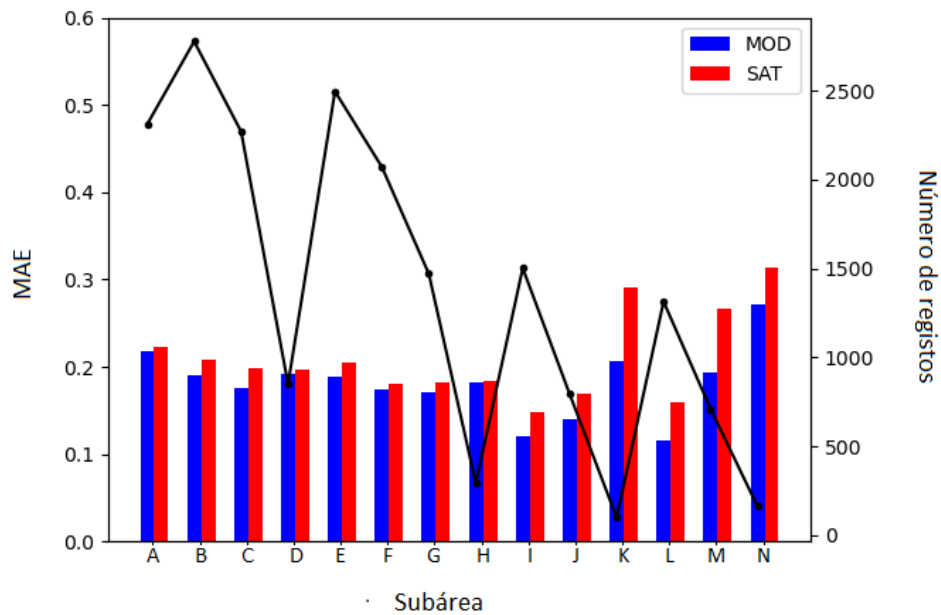


FIGURA 2.7: Valores dos MAE e número de registos por cada subárea definida no capítulo 1.

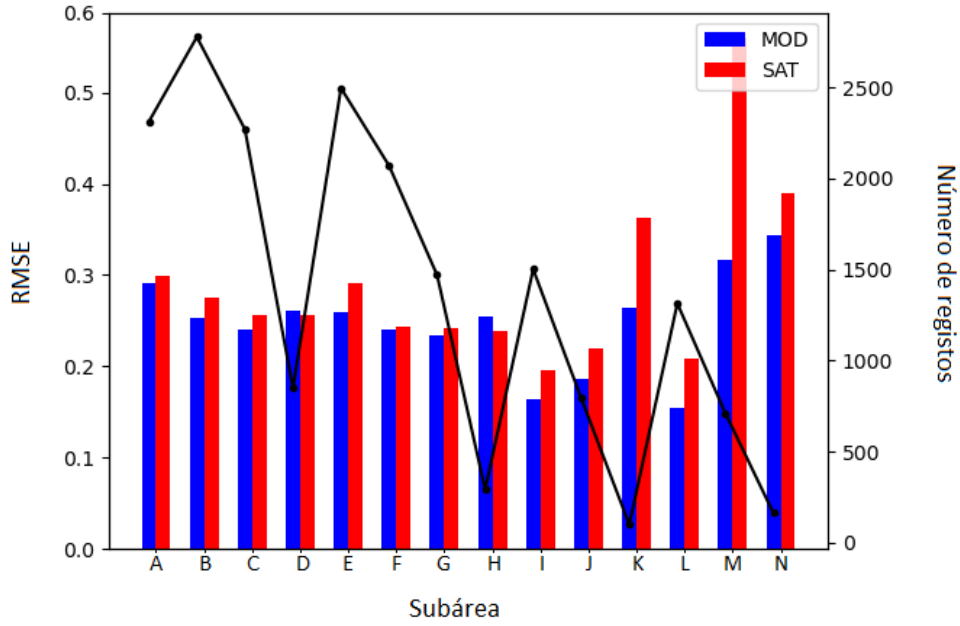


FIGURA 2.8: Valores dos RMSE e número de registros por cada subárea definida no capítulo 1.

A diferença entre dados  $TSM_{GDP}$  e  $TSM_{MOD}$  varia entre  $-1,8$  e  $2,6^{\circ}\text{C}$  (amplitude de  $4,3^{\circ}\text{C}$ ) enquanto a diferença entre  $TSM_{GDP}$  e  $TSM_{SAT}$  varia entre  $-1,8$  e  $4^{\circ}\text{C}$  (amplitude de  $5,8^{\circ}\text{C}$ ) (Figura 2.9). Ambos os histogramas se encontram centrados em zero e o com mais percentagem de diferenças negativas do que positivas.

Através da análise do gráfico de dispersão (Figura 2.10) de valores dos erros absolutos AE por cada distância entre posições dos dados GDP e os dados de previsão (MOD e SAT), pode-se concluir que uma maior distância não implicará um maior erro. Embora uma maior distância geográfica entre a posição (latitude e longitude) onde foi registada a  $TSM_{GDP}$  e a posição na malha de  $TSM_{MOD}$  e  $TSM_{SAT}$ , não havendo evidências de que haja uma correlação entre os dois, os resultados sugerem que as  $TSM_{SAT}$  são mais afetadas pela aplicação do método dos vizinhos mais próximos (Não se verificam erros superiores a  $2^{\circ}\text{C}$  nos dados MOD com distâncias superiores a  $0,1^{\circ}$  e verifica-se nos SAT, não necessariamente relacionado com a distância entre pontos) (Figura 2.11).

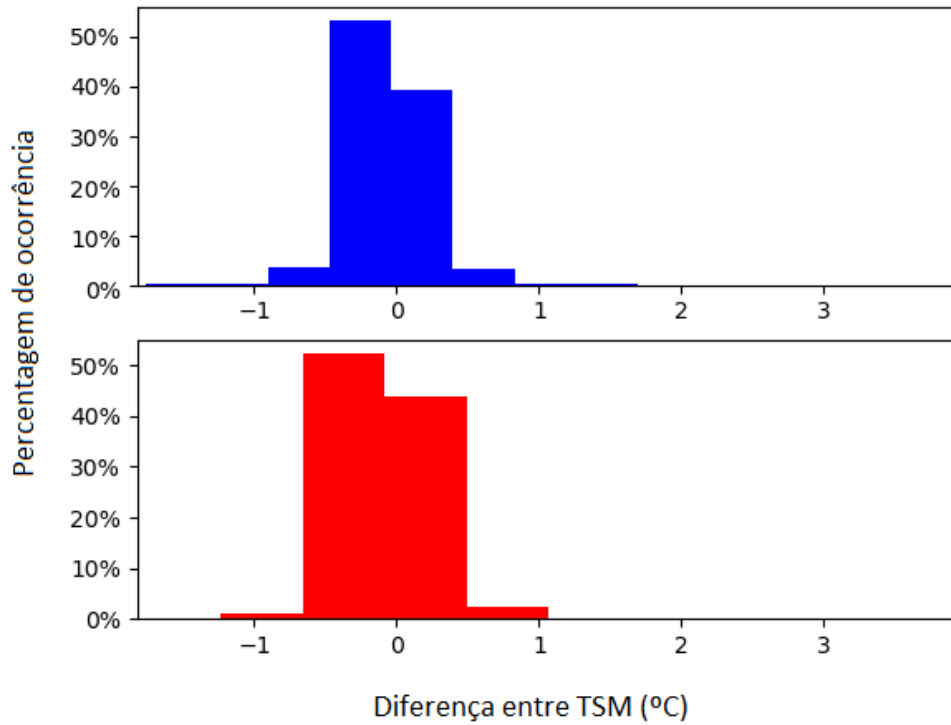


FIGURA 2.9: Histograma das diferenças entre TSMGDP e TSMMOD (Azul) e TSMGDP e TSMSAT (Vermelho) em percentagem.

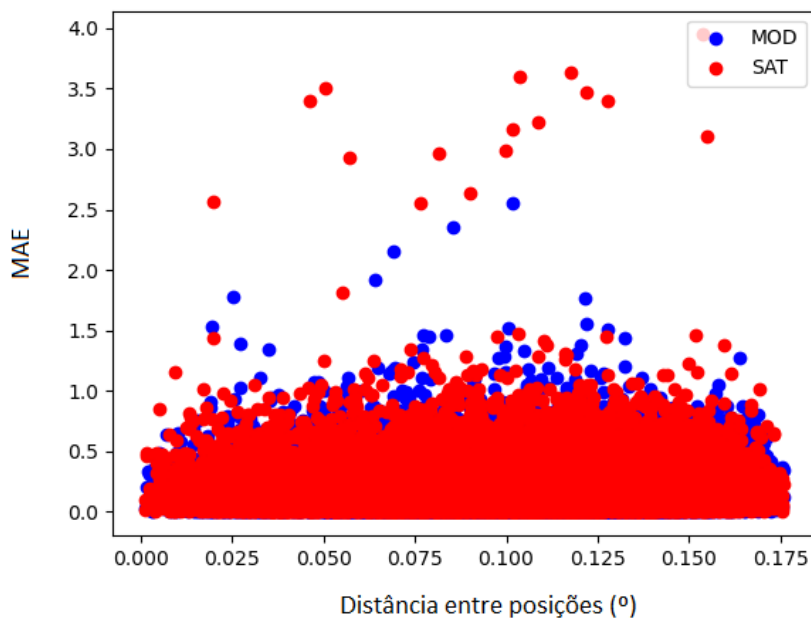


FIGURA 2.10: Correlação entre MAE e distância entre posições das boias durante o registo e a posição do registo MOD (Azul) ou SAT (vermelho) na malha.

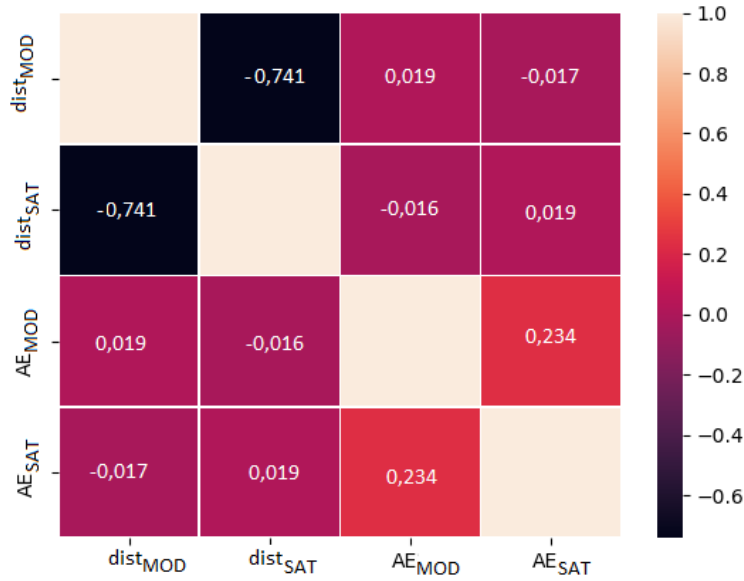


FIGURA 2.11: Matriz de correlação entre os valores de TSM e distância entre posições das boias durante o registo e a posição do registo MOD e SAT na malha.

## 2.3 Análise das boias lançadas pelo NRP Sagres

Esta secção tem o objetivo de descrever detalhadamente a análise efetuada ao comportamento das boias que foram lançadas pelo NRP Sagres em 2020 entre 6 de janeiro de 2020 e 22 março. O navio lançou 16 boias entre o sul da costa de Portugal Continental e o oeste de África do Sul, sendo que 5 delas se encontravam na AI durante o período em estudo. Estas boias possuem os seguintes números de identificação do *World Meteorological Organization* (WMO): 1301592, 6202656, 6202657, 1301589, 1301590, sendo que 2 boias (6202657 e 1301589) cobriram as áreas I, J, K, L, M, N enquanto as restantes 3, as áreas C, D, G, H (Figura 2.12). Os dados de TSM recolhidos ao longo das trajetórias destas 5 boias, durante o período de estudo, foram analisados e foram identificadas as diferenças entre os três conjuntos de dados (GDP, MOD e SAT).

A boia 1301592 tem o primeiro registo no dia 9 de janeiro de 2020, na posição 31.9°N e 14.2°W, e o último no dia 29 de junho de 2020, na posição 28.1°N e 18.3°W. O seu trajeto ocorreu sempre na área G, no qual foram identificados quatro movimentos giratórios. Uma destas situações foi destacada e denominada de A que ocorreu entre janeiro e fevereiro de 2020 (situação que será estudada adiante). Nesta

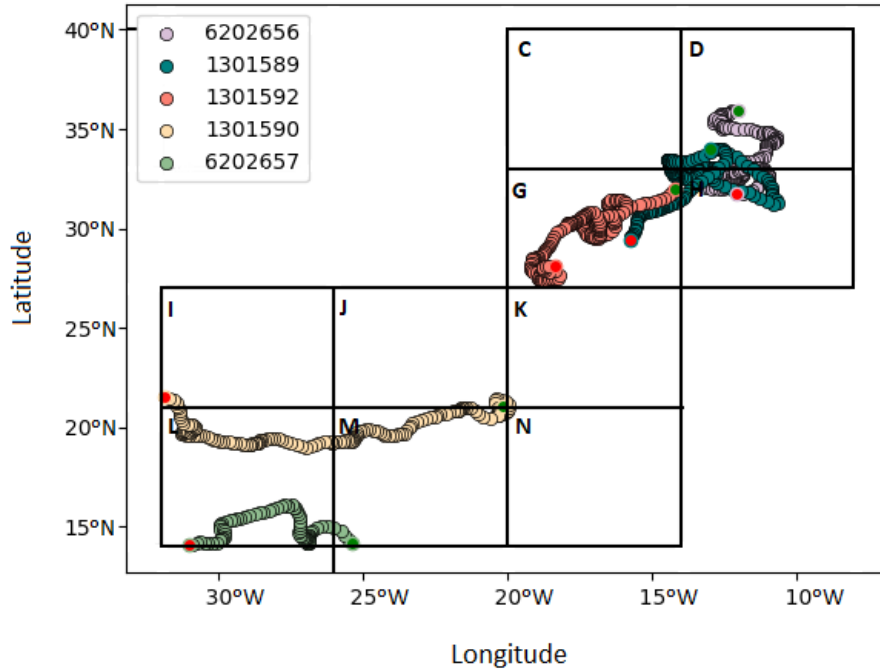


FIGURA 2.12: Posições geográficas das boias lançadas pelo NRP Sagres na AI durante o período de estudo.

situação, a boia movimentou-se entre o arquipélago da Madeira e das Canárias, e teve tendência a deslocar-se para sudoeste tendo passado por um *eddy*, que a fez derivar para norte efetuando um círculo e depois novamente para sudoeste (Figura 2.13). Durante a trajetória desta boia, a diferença entre os valores de  $TSM_{GDP}$  e os  $TSM_{MOD}$  e  $TSM_{SAT}$ , ambas variaram entre  $-0,6$  e  $0,5^{\circ}\text{C}$ , respetivamente, com uma média aproximadamente igual a  $-0,1^{\circ}\text{C}$ . Em valores médios, o  $RMSE_{MOD}$  foi superior à  $RMSE_{SAT}$  com os respetivos valores de  $0,22$  e  $0,21$ . O mesmo se verifica com o  $MAE_{MOD}$  que também é superior à  $MAE_{SAT}$  com  $0,17$  e  $0,15$ . O  $MBE_{MOD}$  é igual ao  $MAE_{SAT}$  com o valor de  $-0,1$ . No entanto, o coeficiente de Kendall é de igual valor. Durante os 12 meses, o viés de ambos os conjuntos de dados SAT e MOD foi tanto negativo como positivo, sendo tendencialmente positivo até meio de março (sobre-estimação) e os restantes meses que se seguiram tendencialmente negativo (subestimação). Já o MAE apresentou mais picos de erros no final do mês de janeiro, início do mês de abril, início do mês de maio e final do mês de junho, onde as distâncias entre as boias e as malhas de dados MOD e SAT verificam-se maiores (Figura 2.14 e figura 2.15).



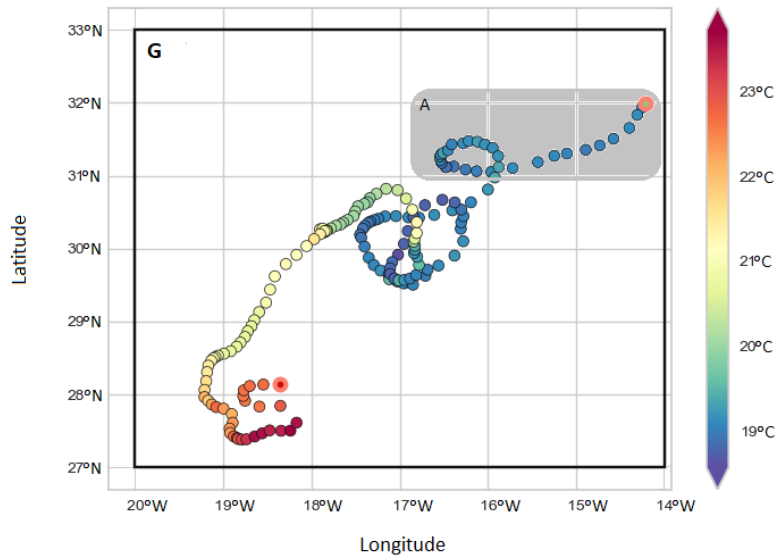


FIGURA 2.13: Posições geográficas da boia 1301592, entre 9 de janeiro 2020 e 29 de junho 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identifica-se com a letra A, uma situação onde os valores previstos da TSMMOD e TSM-SAT registaram maiores diferenças em relação ao TSMGDP.

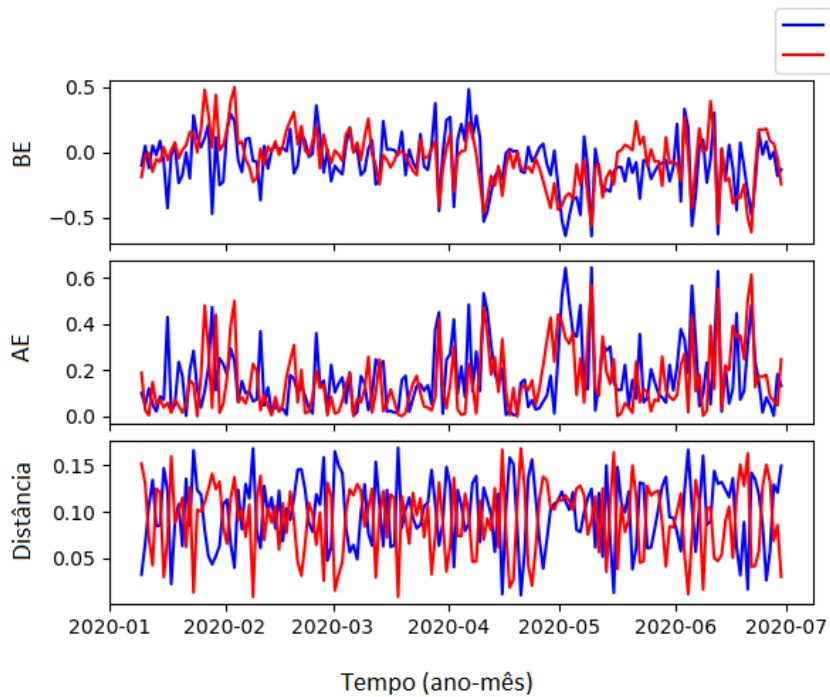


FIGURA 2.14: Variação temporal do BE, AE e distância entre a posição geográfica da boia e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD.

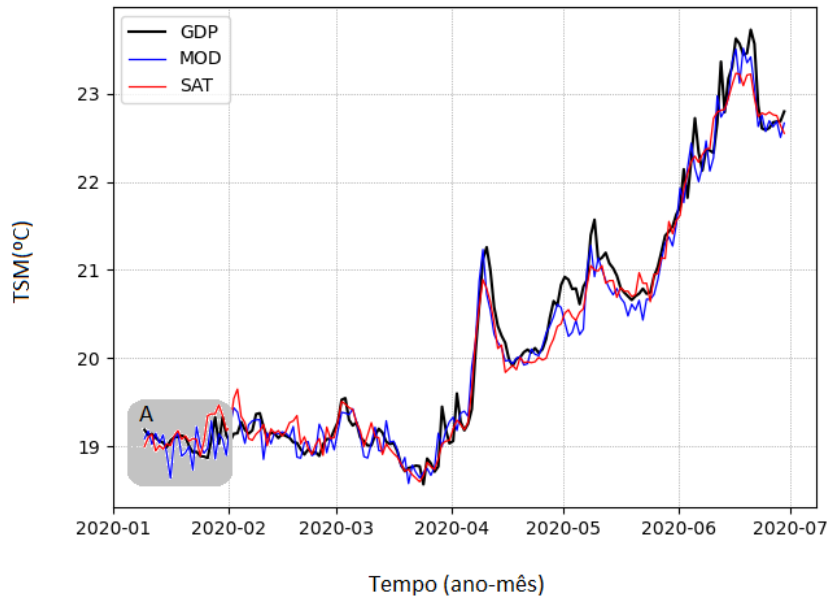


FIGURA 2.15: Variação temporal da TSM (azul:  $TSM_{GDP}$ , laranja:  $TSM_{MOD}$ ; verde:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301592 durante o período analisado (09 de janeiro a 29 de junho 2020). O círculo a cinzento claro identifica o período correspondente à situação A.

Analisando as diferenças entre  $TSM_{MOD}$  e  $TSM_{SAT}$  em relação a  $TSM_{GDP}$  durante a situação A, verificam-se temperaturas semelhantes em magnitude e variabilidade no início do trajeto, apesar do produto MOD ter apresentado alguns mínimos locais acentuados. No entanto, quando a boia começa a efetuar o movimento circulatório identificado na Figura 2.17, verificou-se que, tanto o SAT como MOD, apresentaram maiores diferenças em relação às temperaturas registadas *in-situ*. O SAT apresentou irregularidades na primeira descida de temperaturas, com variações pouco bruscas enquanto o MOD teve variações de sobre e subestimações sucessivas, sendo as mais bruscas as no dia 16 de janeiro, onde previu uma descida de temperatura  $0.3^{\circ}\text{C}$  abaixo da  $TSM_{GDP}$  (Figura 2.16). A boia 6202656, que tem como primeiro registo o dia 9 de janeiro de 2020, na posição  $33.9^{\circ}\text{N}$  e  $12.9^{\circ}\text{W}$  e o último registo no dia 29 de junho de 2020, na posição  $29.4^{\circ}\text{N}$  e  $15.7^{\circ}\text{W}$ , percorreu as áreas D e H entre noroeste da Madeira e norte das Canárias, em direção a sul, tendo sido identificados quatro movimentos giratórios (Figura 2.17).

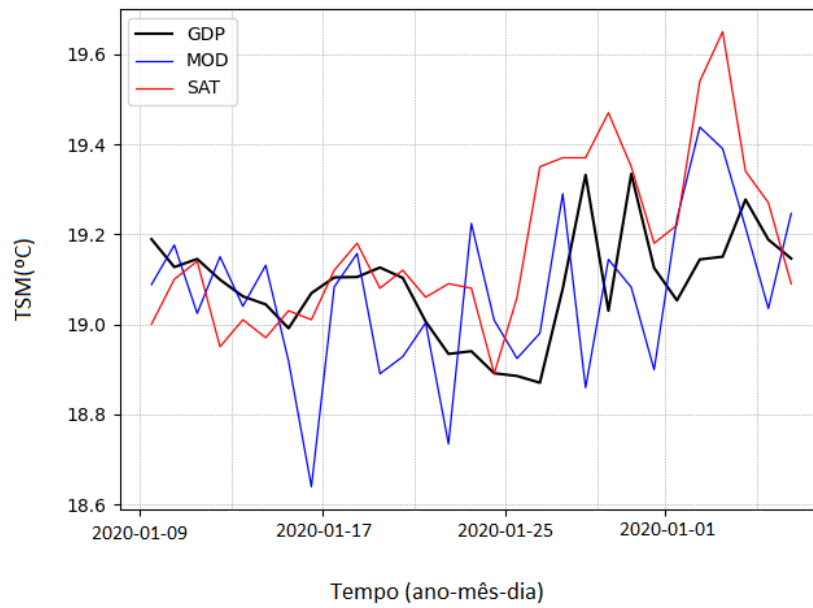


FIGURA 2.16: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301592 durante situação A.

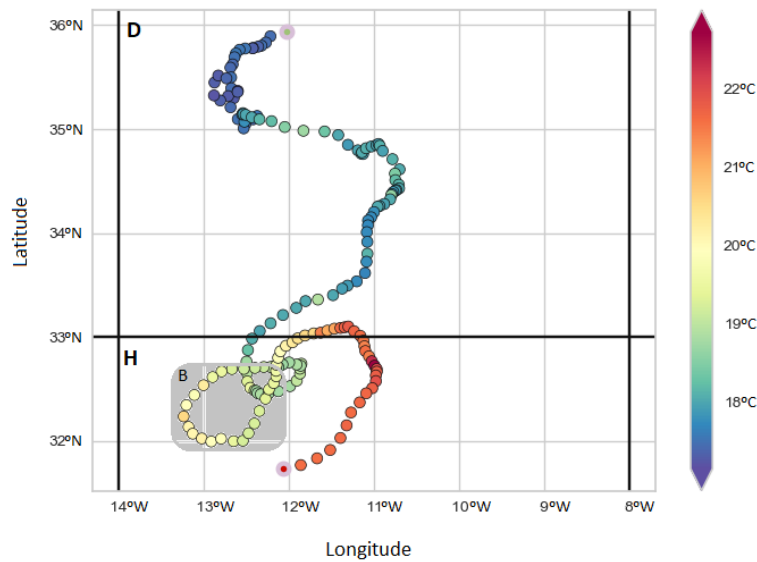


FIGURA 2.17: Posições geográficas da boia 6202656, entre 9 de janeiro 2020 e 29 de junho 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identifica-se com a letra B, uma situação onde os valores previstos da  $TSM_{MOD}$  e  $TSM_{SAT}$  registaram maiores diferenças em relação ao  $TSM_{GDP}$ .

Para a trajetória considerada, entre 9 de janeiro e 29 de junho de 2020, a diferença entre  $TSM_{GDP}$  e  $TSM_{MOD}$  está compreendida entre  $-0,9$  e  $0,8^{\circ}\text{C}$ , em média  $-0,02^{\circ}\text{C}$ , e a diferença entre o primeiro e a TSM SAT foi entre  $-0,9$  e  $0,4^{\circ}\text{C}$ , em média  $-0,01^{\circ}\text{C}$ . A  $RMSE_{SAT}$  é superior à  $RMSE_{MOD}$  com os valores de  $0,24$  e  $0,23$ . Em ambos os métodos, o MAE e o coeficiente de Kendall tem igual valor embora o  $MBE_{SAT}$  tem o valor de  $-0,1$  e  $MBE_{MOD}$  de  $-0,02$  (Figura 2.18). Durante

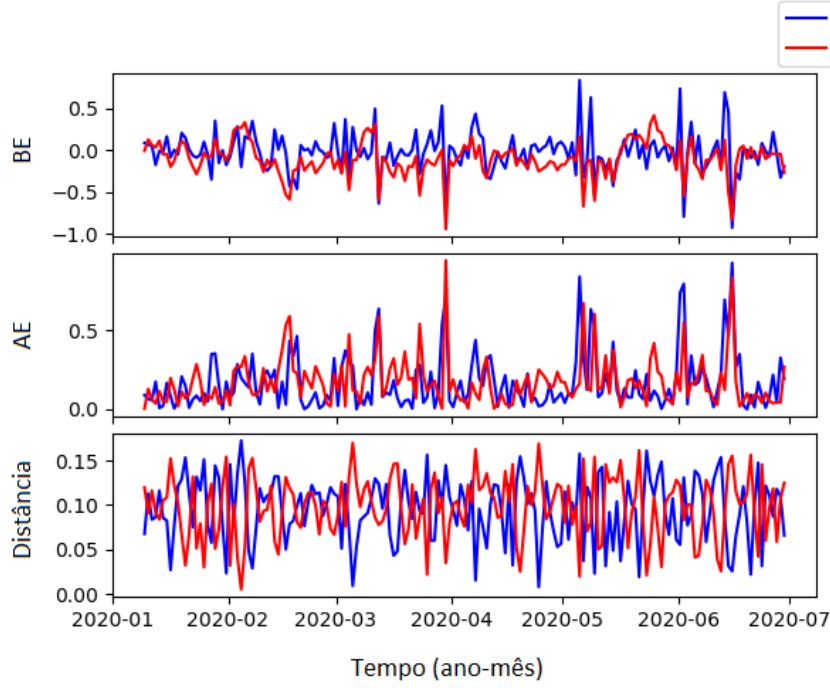


FIGURA 2.18: Variação temporal do BE, AE e distância entre a posição geográfica da boia 6202656 e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD, entre 9 de janeiro 2020 e 29 de junho 2020.

a trajetória da boia 6202656, destaca-se a situação B, percorrida entre maio e junho de 2020, quando a boia derivou tendencialmente para sul com algumas oscilações de longitude e encontrou um *eddy*, identificado através da trajetória da boia que sofreu um movimento circulatório no sentido ciclónico, registando temperaturas entre os  $19$  e os  $21^{\circ}\text{C}$ . Em detalhe, no início da situação B, verifica-se um deslocamento no tempo nos registos  $TSM_{MOD}$ , indicando que, de alguma forma, o MOD antecipou as rápidas alterações de temperatura medidas *in-situ*, enquanto o produto SAT não registou estas alterações, mantendo valores mais constantes e sistematicamente inferiores aos valores *in-situ* (Figura 2.19 e 2.20).

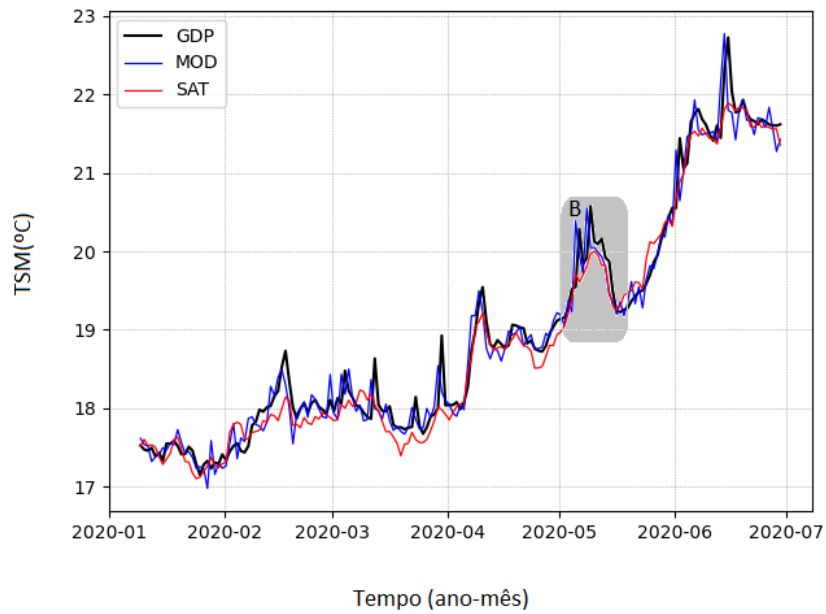


FIGURA 2.19: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 6202656 durante o período analisado (09 de janeiro a 29 de junho 2020). O círculo a cinzento claro identifica o período correspondente à situação B.

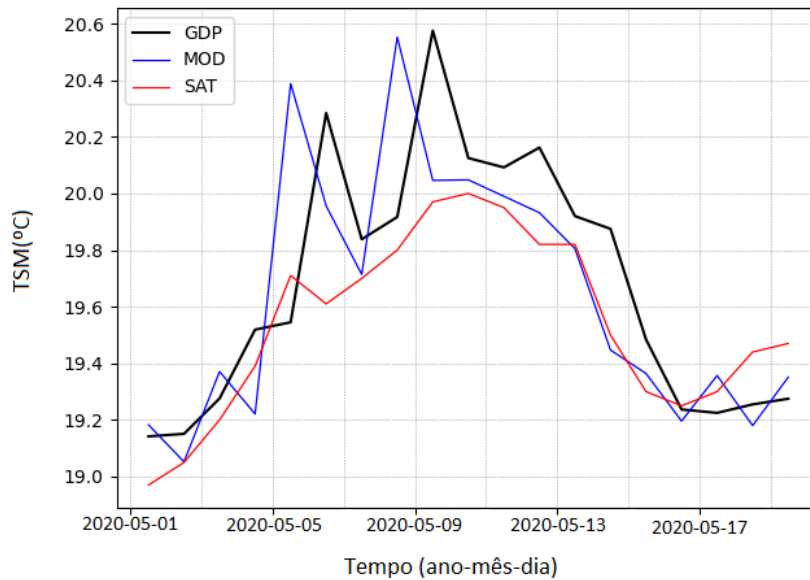


FIGURA 2.20: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 6202656 durante situação B.

Dentro da AI, a boia 6202657 apresentou 87 registos de TSM, medidos entre 1 de abril de 2020, na posição 14.2°N 25.4°W e 26 de junho de 2020, na posição 14.1°N 31°W. O seu trajeto ocorreu entre as áreas L e M, com oscilações de direção norte-sul e efetuou apenas um movimento giratório (Figura 2.21).

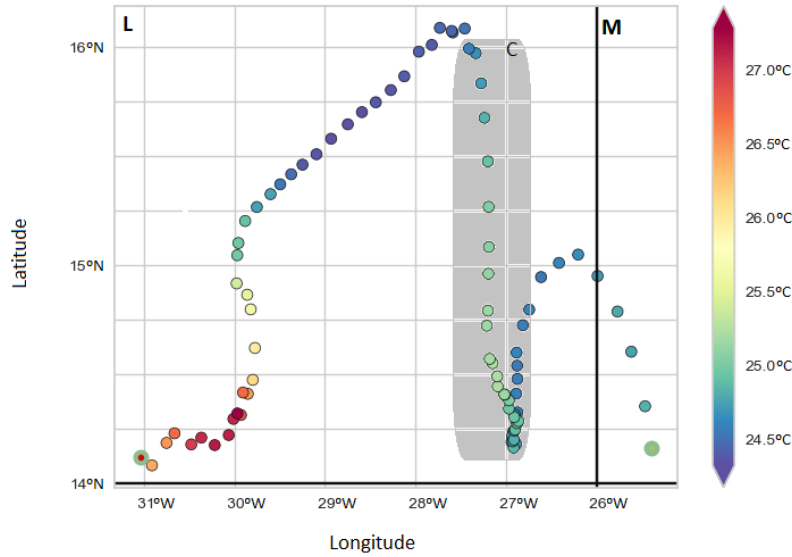


FIGURA 2.21: Posições geográficas da boia 6202657, entre 9 de janeiro 2020 e 29 de junho 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identifica-se com a letra C, uma situação onde os valores previstos da TSMMOD e TSM-SAT registaram maiores diferenças em relação ao TSMGDP.

A diferença entre  $TSM_{GDP}$  e  $TSM_{MOD}$  está compreendida entre  $-0.6^{\circ}$  e  $0.3^{\circ}\text{C}$  (média de  $-0.1^{\circ}\text{C}$ ) e a diferença entre o primeiro e a  $TSM_{SAT}$  foi entre  $-0.8^{\circ}$  e  $0.4^{\circ}\text{C}$  (média de  $-0.2^{\circ}\text{C}$ ). A  $RMSE_{MOD}$  é inferior à  $RMSE_{SAT}$  com os valores de 0,16 e 0,27, respetivamente. Também a  $MAE_{MOD}$  é inferior à  $MAE_{SAT}$  com 0,12 e 0,20. O coeficiente de Kendall é superior no MOD com o valor de 0,98 enquanto este coeficiente na SAT é 0,96. No geral, esta boia ao longo do tempo é bem acompanhada pelo SAT e MOD tendo estes pouco viés, à exceção da situação C, centrada no dia 1 de maio de 2020, onde ocorreu um erro elevado nos SAT que não se verifica no MOD, não relacionado com a distância entre a posição da boia e as malhas (Figura 2.22).

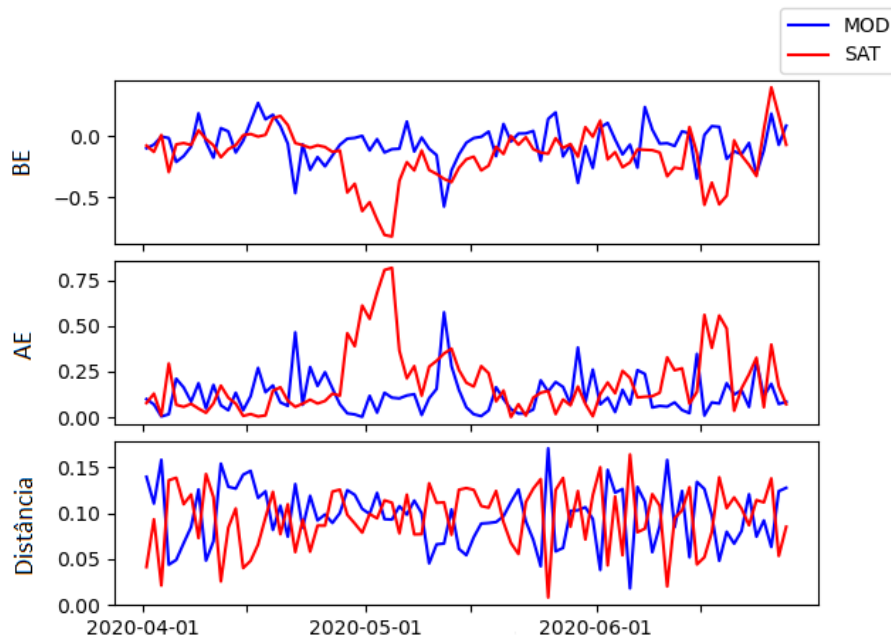


FIGURA 2.22: Variação temporal do BE, AE e distância entre a posição geográfica da boia 6202657 e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD.

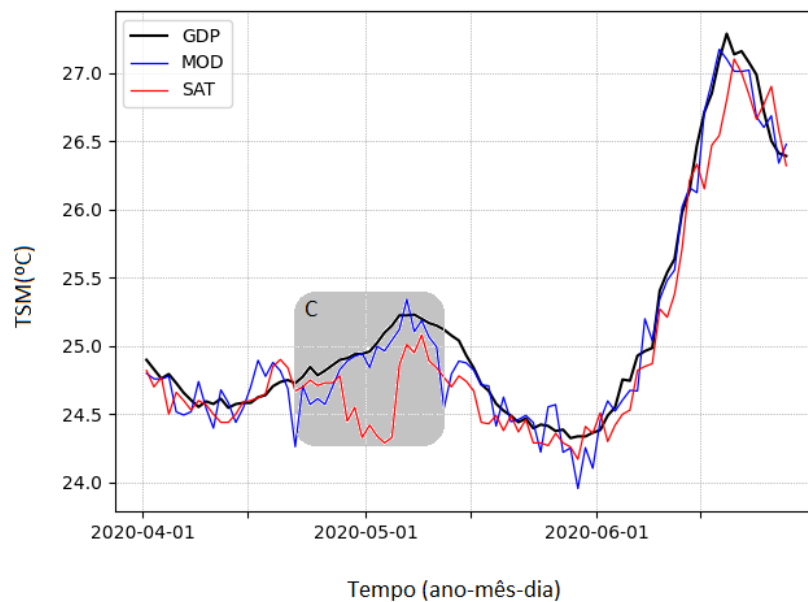


FIGURA 2.23: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 6202657 durante o período analisado (1 a 15 de janeiro 2020). O círculo a cinzento claro identifica o período correspondente à situação C.

Durante o trajeto da boia 6202657 na AI, foi identificada a situação C, que ocorreu no período entre 1 e 15 de janeiro de 2020, quando o deslocamento da boia estava com tendência para sul e ocorreu uma inversão deste deslocamento para norte. Durante a situação C, a boia registou um aumento gradual da temperatura, atingindo valores superiores a  $25^{\circ}\text{C}$ , aumento este que foi também reproduzido pelo MOD. No entanto, observando a Figura 2.23, verifica-se que a  $TSM_{SAT}$  não acompanhou esta variação, registando, pelo contrário, uma diminuição abrupta dos valores de TSM, entre 28 de abril de 2020 e 05 de maio de 2020, não concordantes com os valores registados *in-situ* pela boia (Figura 2.24). A boia 1301589, com 174 registos entre 8 de janeiro de 2020, na posição  $34^{\circ}\text{N}$   $13.1^{\circ}\text{W}$  e 29 de junho de 2020, na posição  $29.4^{\circ}\text{N}$   $15.7^{\circ}\text{W}$ . Esta percorreu 4 áreas: C, D, G e H e na sua trajetória foram identificados 4 movimentos giratórios (Figura 2.25).

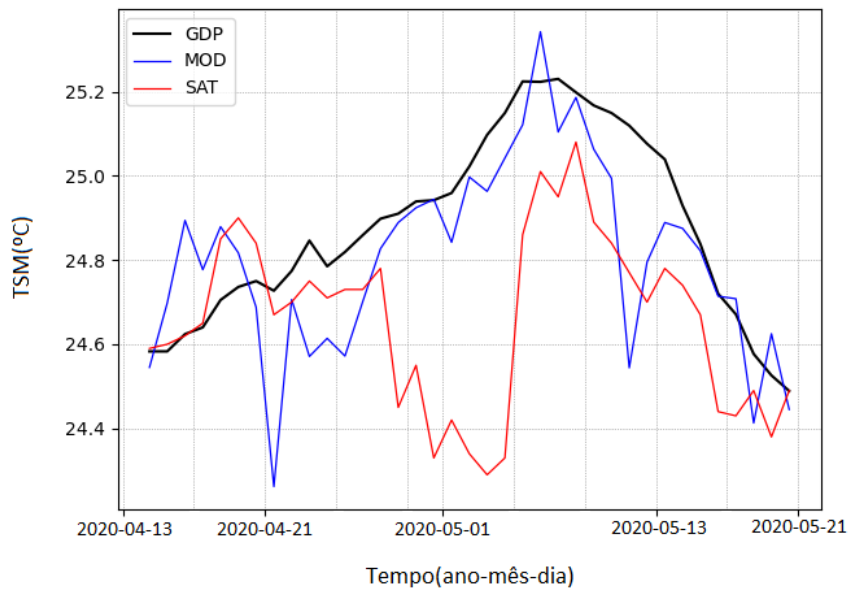


FIGURA 2.24: Variação temporal da TSM preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 6202657 durante a situação C.



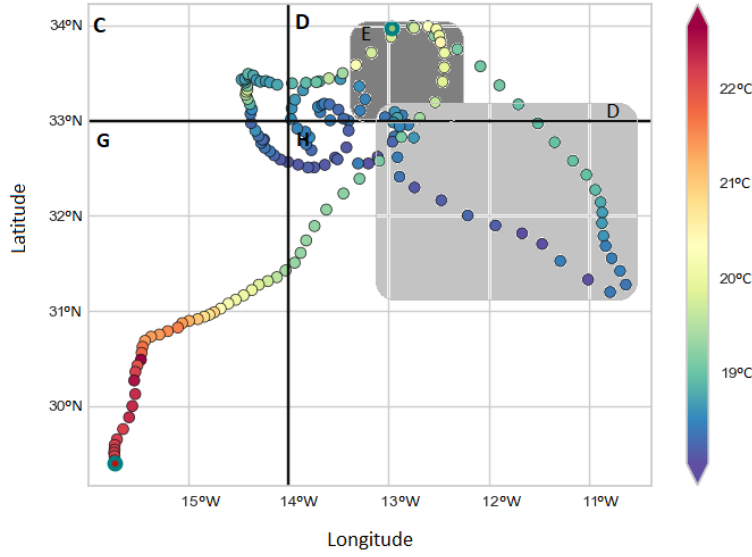


FIGURA 2.25: Posições geográficas da boia 1301589, entre 1 e 15 janeiro 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identificam-se com duas letras, E e D, situações onde os valores previstos da TSMMOD e TSMSAT registaram maiores diferenças em relação ao TSMGDP.

A diferença entre  $TSM_{GDP}$  e  $TSM_{MOD}$  esta compreendida entre  $0.8^\circ$  e  $1.2^\circ\text{C}$  (média  $-0.05^\circ\text{C}$ ) e a diferença entre o primeiro e a  $TSM_{SAT}$  foi entre  $-0.96^\circ$  e  $0.58^\circ\text{C}$  (média  $-0.07^\circ\text{C}$ ). A  $RMSE_{MOD}$  é superior à  $RMSE_{SAT}$  com os respetivos valores de 0,227 e 0,233 e a  $MAE_{MOD}$  é inferior à  $MAE_{SAT}$  com 0,161 e 0,174. No entanto o coeficiente de Kendall é superior na  $TSM_{SAT}$  com o valor de 0,985 enquanto este coeficiente na  $TSM_{MOD}$  é 0,985 (Figura 2.26). Para a análise, são realçadas 2 situações distintas. Na situação D, entre fevereiro e março de 2020, esta boia derivante executou uma inversão rápida de movimento estando primeiramente com tendência a movimentar-se para sudeste e posteriormente para noroeste, verificando-se um aumento e posterior diminuição de temperatura, com valores entre os  $18^\circ$  e  $19^\circ\text{C}$  (Figura 2.27).

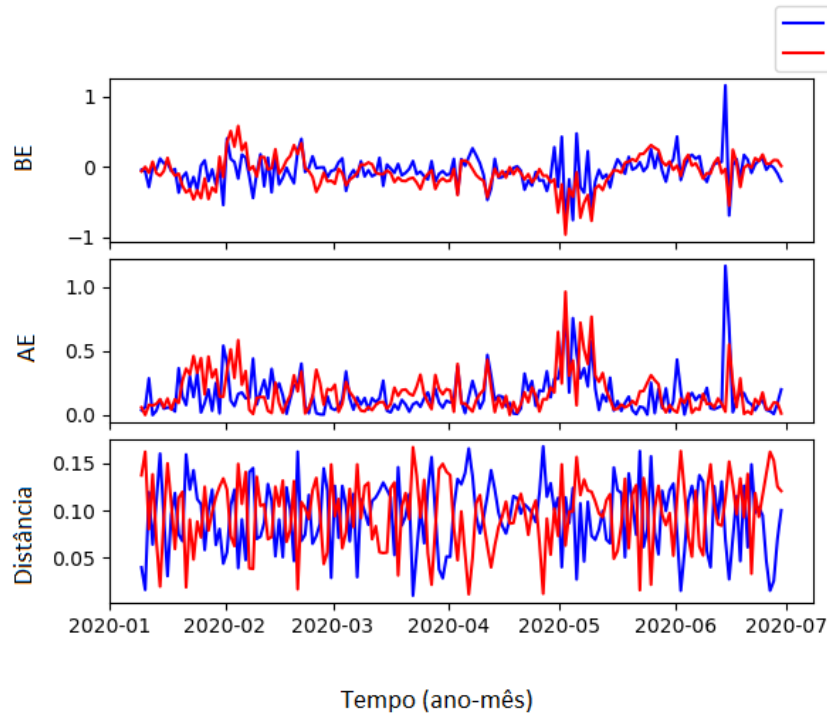


FIGURA 2.26: Variação temporal do BE, AE e distância entre a posição geográfica da boia 1301589 e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD.

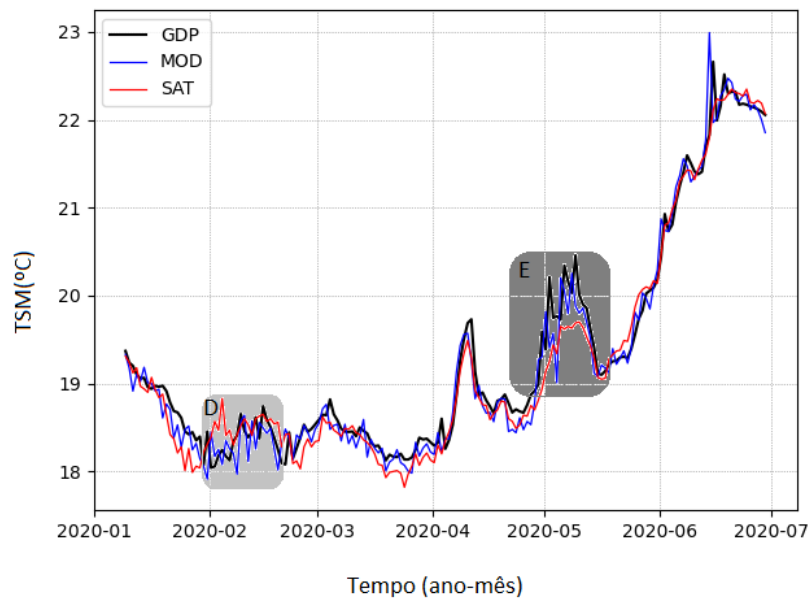


FIGURA 2.27: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301589 durante o período analisado. A sombreado identifica-se o período correspondente às situações D e E.

Durante este período, a  $TSM_{SAT}$  apresentou um máximo no dia 4 de fevereiro de 2020, que não está concordante com GDP. A  $TSM_{MOD}$ , como referido no caso anterior, teve muitas oscilações de temperatura com picos de temperatura bastante inferiores aos medidos *in-situ* (Figura 2.28).

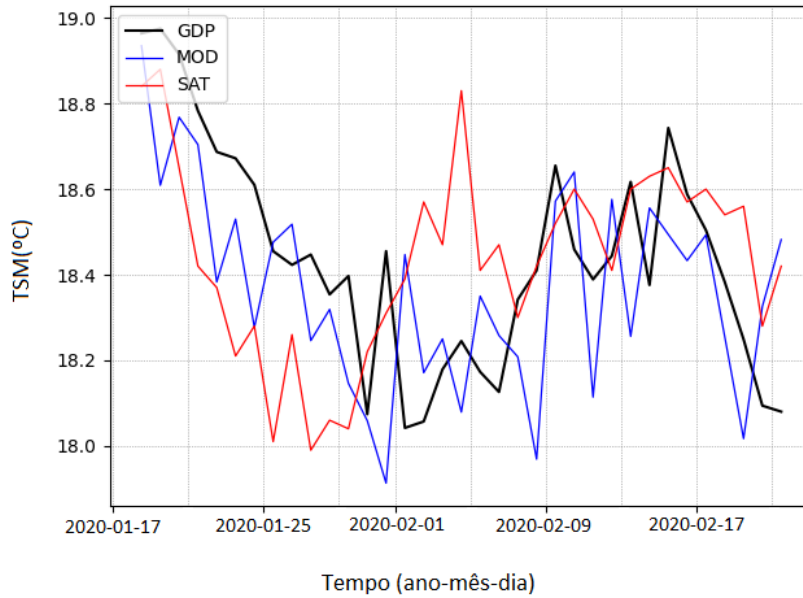


FIGURA 2.28: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 61301589 durante a situação D.

Já na situação E, na primeira quinzena do mês de maio de 2020, a boia aparentemente derivou com um *eddy*, visível através das várias rotações que executou, e tendo permanecido durante algum tempo na mesma região. Durante este período, verificou-se uma grande oscilação nos valores medidos *in-situ* com três picos de temperatura. O MOD conseguiu representar dois desses picos, mas um deles com diminuição de temperatura em vez de aumento. A  $TSM_{SAT}$  não demonstrou picos tão elevados e manteve-se com valores inferiores, mas mais constantes e graduais (Figura 2.29).

Por último, a boia 1301590 apresentou 163 registos, sendo o primeiro a 17 de janeiro de 2020 na posição 21.1°N 20.2°W e o último registo analisado foi de 27 de junho de 2020, na posição 21.5°N 31.9°W (Figura 2.30).

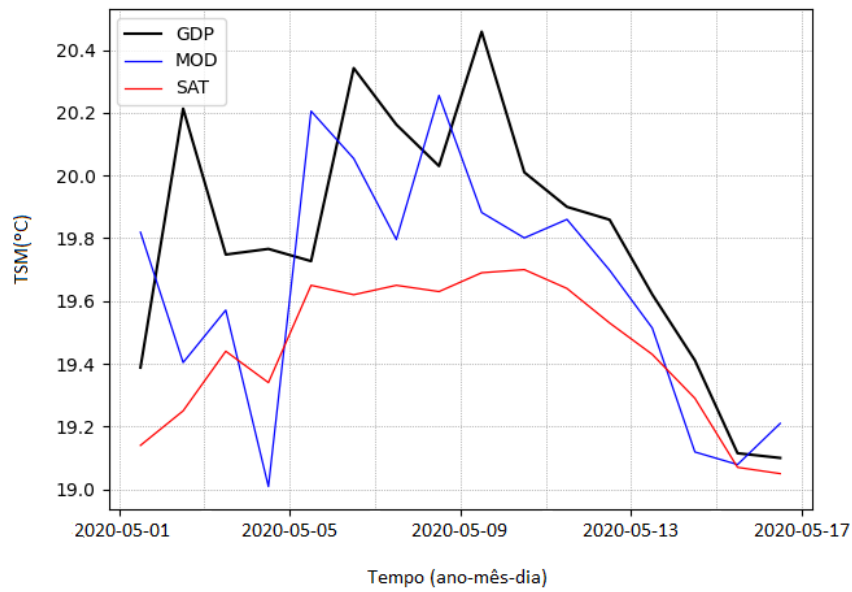


FIGURA 2.29: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 61301589 durante a situação E.

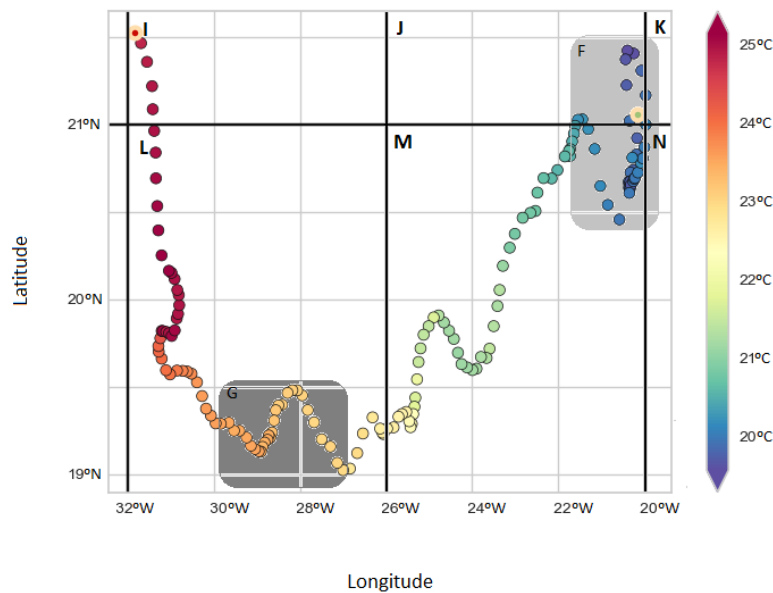


FIGURA 2.30: Posições geográficas da boia 1301590, entre 17 de janeiro e 27 de junho de 2020. A primeira e última posição estão marcadas por um círculo laranja. A sombreado identifica-se com as letras F e G, situações onde os valores previstos da  $TSM_{MOD}$  e  $TSMSAT$  registaram maiores diferenças em relação ao  $TSM_{GDP}$ .

A diferença entre  $TSM_{GDP}$  e  $TSM_{MOD}$  está compreendida entre  $-0,6$  e  $0,5^{\circ}\text{C}$ , numa média de  $-0,1^{\circ}\text{C}$ , e a diferença entre o primeiro e a  $TSM_{SAT}$  foi entre  $-0,6$  e  $0,2^{\circ}\text{C}$ , em média  $-0,1^{\circ}\text{C}$ . Em valores médios, o  $RMSE_{MOD}$  tem igual valor de  $0,17$  que o  $RMSE_{SAT}$ . Os MAE são igual com o valor de  $0,13$ . Também o coeficiente de Kendall tem igual valor e o  $MBE_{SAT}$  é superior com um valor de  $-0,11$  enquanto  $MBE_{MOD}$  tem um valor de  $-0,06$ . Na situação F verifica-se um pico de viés positivo e a situação G um pico de viés negativo (Figura 2.31).

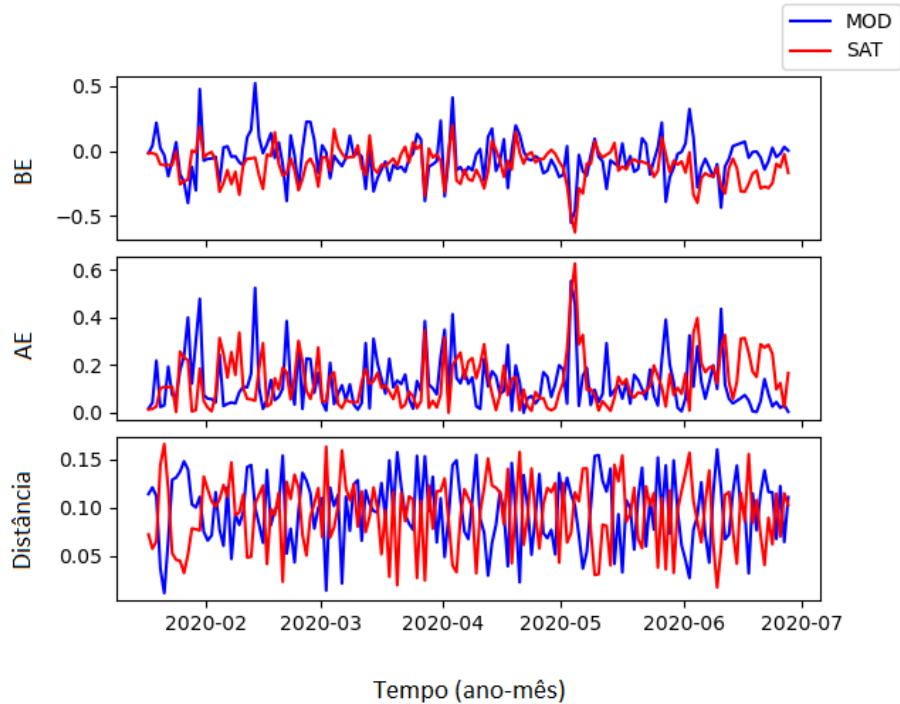


FIGURA 2.31: Variação temporal do BE, AE e distância entre a posição geográfica da boia 1301590 e a posição nas malhas onde foi retirado o registo dos dados SAT e MOD.

Durante o percurso da boia 1301590 foi identificada a situação F, ocorrida na primeira metade do mês de fevereiro de 2020 (Figura 2.32). A boia esteve possivelmente a derivar num *eddy* tendo efetuado uma rotação no sentido anticiclónico com uma posterior tendência de deriva para oeste, com algumas oscilações de latitude. Durante este período, a boia registou um pico mínimo de temperatura que foi acompanhado pelas outras duas fontes MOD e SAT (Figura 2.33). A  $TSM_{MOD}$  representou bem esta diminuição, mas posteriormente antecipou o aumento de temperatura registado pela boia, enquanto  $TSM_{SAT}$  apresentou antecipadamente a diminuição de temperatura e um atraso no seu posterior aumento.

Na situação G, observa-se uma subida gradual de temperatura *in-situ* mas,

nos primeiros dias do mês de maio, tanto o SAT como o MOD detetam uma diminuição brusca de temperatura (Figura 2.34).

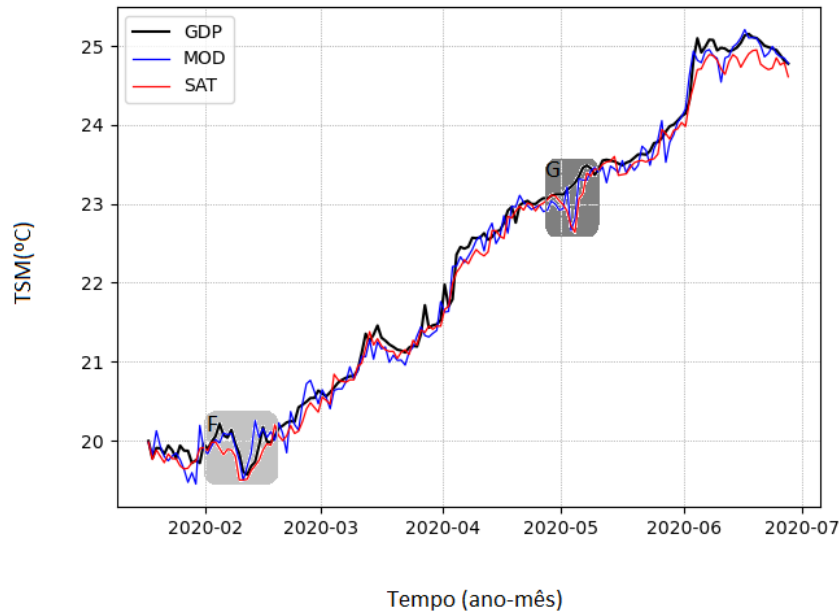


FIGURA 2.32: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301590 durante o período analisado (17 de janeiro a 27 de junho 2020). O círculo a cinzento claro identifica o período correspondente à situação F e o cinzento escuro identifica o correspondente à situação G.

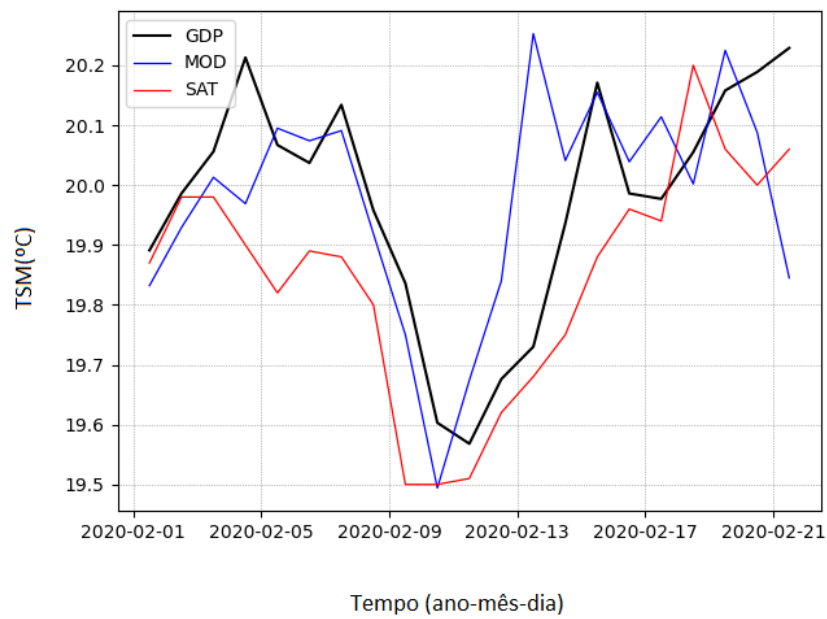


FIGURA 2.33: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301590 durante a situação F.



FIGURA 2.34: Variação temporal da TSM (preto:  $TSM_{GDP}$ , azul:  $TSM_{MOD}$ ; vermelho:  $TSM_{SAT}$ ) correspondente ao trajeto da boia 1301590 durante a situação G.





## Capítulo 3

### Discussão dos Resultados

Após a remoção dos registos da boia 4101718 que apresentava incongruências com ambos os conjuntos de dados MOD e SAT, a diferença entre  $TSM_{GDP}$  e  $TSM_{MOD}$  está compreendida entre  $-1.8^\circ$  e  $2.6^\circ\text{C}$ , numa média de  $-0.1^\circ\text{C}$ , e a diferença entre o primeiro e a  $TSM_{SAT}$  foi entre  $-1.8^\circ$  e  $4.0^\circ\text{C}$ , em média  $-0.1^\circ\text{C}$ . Ou seja, apesar de, em média, as diferenças entre os dados serem de igual valor, as  $TSM_{SAT}$  tem maior dispersão. Deste modo, verificou-se que as  $TSM_{SAT}$  têm maior oscilação tanto para erros positivos como para negativos. Verificou-se também que, quando as boias têm um movimento constante na mesma direção, as TSM dos três conjuntos têm erros de menor ordem do que aquelas que tiveram perante fenómenos de mesoescala.

A  $TSM_{MOD}$  apresentam maior valor de  $\sigma$ , o que indica que as  $TSM_{SAT}$  apresentam melhor precisão nas medições. O RMSE e o MAE foi superior na  $TSM_{SAT}$ , pelo que se verifica que o seu conjunto de dados é aquele que no total possui mais diferenças entre os valores. O MBE foi superior na  $TSM_{SAT}$  o que indica que, apesar de ambos os métodos terem subestimado os valores de  $TSM_{GDP}$  (Tanto geral como por subárea), a  $TSM_{SAT}$  efetua-o com maior amplitude (Tabela 2.2).

Uma vez comparados os erros entre as subáreas e a média global dos erros, apurou-se que o MBE do MOD é superior à média global, em valor absoluto, nas áreas A, E, G, K e N enquanto o SAT é nas áreas B, C, D, F, I, J, K, L, N. As áreas A e E são as únicas em que o MOD subestima mais a realidade do que o SAT. Quanto ao MAE do MOD este é superior à média nas subáreas A, B, D, E, K, M, N enquanto o SAT é superior à media nas A, B, E, K, M, N (A subárea D tem um MAE MOD superior ao SAT). O RMSE do MOD é superior à média nas subáreas A, D, E, K, M, N enquanto o SAT é nas subáreas E, K, M e N (As subárea A e D tem um RMSE MOD superior ao SAT). Por outras palavras, as subáreas E, K, M

e N tiveram precisões abaixo da média geral nos dois métodos enquanto a subárea A teve abaixo da média geral apenas no MOD.

No entanto, comparando cada um dos erros MOD com os SAT, conclui-se que em termos de precisão (MAE e RMSE) apenas na subárea H o SAT foi mais preciso do que o MOD. Nas subáreas D, F e G, os dois métodos tiveram erros semelhantes. Mas em todas as restantes subáreas, isto é, nas subáreas A, B, C, E, I, J, K, L, M, N o MOD teve melhor precisão.

Posto isto, as áreas com piores previsões são a E, K, M e N com todos os MAE e RMSE superiores às suas médias globais. Também se verifica que na maioria das subáreas os SAT apresentam valores de erros superiores ao MOD.

De modo geral, a TSM foi crescente de julho a setembro de 2019 e de março a julho de 2020. Através dos movimentos de deriva das boias, concluiu-se que a AzC é dominada por fenómenos de mesoescala provocando os movimentos circulares das boias nas áreas A, B, C, E, F e G e que sofre um relaxamento da circulação para Oeste nos meses mais frios e sofre mais influência para Sul nos meses mais quentes. A CaC é a corrente onde se verificou menos variação sazonal e direção constante ao longo do ano. A NEC sofreu uma alteração de latitude para norte nos meses mais quentes e a PC provoca deslocamentos rápidos nas boias. Deste modo, as áreas onde se verificou maior importância no lançamento das boias para obtenção de dados TSM foi ao largo de Portugal Continental, Arquipélago da Madeira, costa noroeste de África e Arquipélago das Canárias devido ao pouco tempo de permanência de boias nestas regiões.

Também se verificou que as subáreas mais próximas do continente são aquelas com menos quantidade de registos (subáreas D, H, K, N). Por outro lado, a região com mais cobertura foi entre 27 e 40°N e 26 e 32°W (subáreas B e E). Também se concluiu que mais boias não implica mais registos pois a movimentação de cada local dita o tempo de permanência da boia em cada subárea. As áreas E, I, J, L e N são aquelas com médias de TSM mais elevadas (superiores a 23°C) enquanto que as mais baixas são a C e D com TSM inferiores a 20°C.

Assim, as maiores diferenças entre valores de TSM surgem quando existem fenómenos de mesoescala. Como demonstrado nas boias lançadas pelo NRP Sagres, quando ocorrem movimentos circulares nas suas trajetórias é quando se verificam alterações nos gráficos de linha de projeções das temperaturas ao longo do tempo, que por vezes seguem os declives da reta das  $TSM_{GDP}$  como na situação F (indicada na figura 2.32), em que apesar de haver diferenças entre valores verifica-se que os

declives das retas são semelhantes e que houve detecção da variação abrupta de TSM, mas também se verifica situações em que não é detetada nenhuma alteração como na situação E, em que o SAT não acompanhou a subida abrupta de temperatura e o MOD detetou uma descida de temperatura.

Em termos de limitações, as boias derivantes apenas têm uma cobertura de 77% e como consequência, a utilização única deste método não é suficiente. Estas têm um tempo de vida médio de 18 meses, o que provoca que todos os anos tenham de ser lançadas aproximadamente 1000 boias para manter a matriz objetivo do programa. Para além disso, uma vez lançadas não podem ser calibradas e estão condicionadas a condições de mar muito adverso. O maior aspeto positivo é medirem os dados *in-situ* que até hoje é a forma de medição que traduz maior confiança e fornece dados em tempo real. Também se verificou que os registos do GDP, uma vez que são medidos no local onde estas se encontram, oferecem pouca quantidade de dados à escala global surgindo a necessidade de recorrer a métodos de interpolação com os erros intrínsecos a este. Por outro lado, verificou-se que 1 das 114 boias possuía uma má calibração durante 349 dias sucessivos, o que manifesta a necessidade de uma utilização conjunta com outros métodos pois mesmo o ground truth do estudo teve, embora em baixa percentagem, um erro associado.

Já o MOD necessita de assimilar dados provenientes de satélites e dados *in-situ* de boias e navios, e nunca poderá ser independente deste outros métodos. No entanto, o aspeto mais positivo é que fornece dados instantâneos a todas as horas do dia, e efetua previsões e análises aos seus registos por forma a mitigar os erros. Para além disso, não tem custos associados (apenas os custos dos modelos em que se apoia) e tem cobertura global.

Por último, os dados SAT, este produto de nível 4 que possui sensores de radiação de infravermelhos e micro-ondas obteve no geral bons resultados em condições mais adversas como nebulosidade. No entanto, verificou-se que a qualidade e precisão dos seus registos era superior nos meses de mais quentes, de maio a agosto, conforme demonstra a Figura 20. Estes carecem de calibração e validação por dados *in-situ*. Neste estudo, verificou-se que em alguns momentos, a atenuação das variações bruscas de temperatura tendo alguma dificuldade em detetar os picos de subida e descida de temperatura.

### 3.1 Boias lançadas pelo NRP Sagres

As boias 1301592, 6202656, 1301589 e 1301590 são aquelas que possuem séries temporais semelhantes, entre janeiro e junho de 2020. A boia 6202657 é a única com registos entre abril e junho de 2020. Geograficamente, as boias 1301590 e a 6202657 encontraram-se nas áreas I, J,K, L, M e N enquanto as restantes nas áreas C, D, G e H. No entanto, verifica-se que as boias 1301592 e 6202656 têm erros muito semelhantes de MAE e RMSE, tal como as boias 301589 e 1301590, embora a boia 1301590 obtenha melhores resultados de erros tanto para o  $TSM_{MOD}$  como para o  $TSM_{SAT}$  (Tabela 3.1).

As boias lançadas pelo NRP Sagres que obtiveram piores resultados foram a 6202656 com piores exatidões MOD (derivou entre latitudes 27 e 40°N e longitudes 8 e 14°W) e a 6202657 piores exatidões SAT (derivou entre latitudes 14 e 21°N e longitudes 20 e 32°W)

	RMSE		MAE		MBE	
	MOD	SAT	MOD	SAT	MOD	SAT
<b>Boia 1301592</b>	0,22	0,21	0,17	0,15	-0,1	
<b>Boia 6202656</b>	0,24	0,23	0,17		-0,02	-0,1
<b>Boia 6202657</b>	0,16	0,27	0,12	0,20	-0,06	-0,16
<b>Boia 1301589</b>	0,23		0,16	0,17	-0,05	-0,08
<b>Boia 1301590</b>	0,17		0,13		-0,06	-0,11

TABELA 3.1: Tabela síntese dos erros das boias lançadas pelo NRP Sagres

# Conclusão

Este estudo permitiu examinar as variações de TSM que ocorreram ao longo de um ano, na região da Macaronésia, bem como analisar os diferentes métodos que as registaram. As diferenças entre os registos de  $TSM_{GDP}$  e as  $TSM_{MOD}$  e  $TSM_{SAT}$  foram analisados e calculados os seus erros associados. Uma vez que os registos GDP são medidos *in-situ*, este conjunto de dados foi considerado a *ground truth* deste estudo, tendo sido comparados os erros das  $TSM_{MOD}$  e  $TSM_{SAT}$  relativamente às  $TSM_{GDP}$ .

Após o tratamento de dados, onde se removeu todos os registos da boia 4101718, foram utilizadas várias medidas de erro (RMSE, MBE, MAE e correlação) para avaliar a diferença entre os três conjuntos de dados diferentes. Comparou-se estes erros por subárea e por mês por forma a estudar a variabilidade de TSM espaciotemporal na região da Macaronésia. Dadas as características de cada metodologia, considerou-se os dados GDP como a *ground truth*.

Em relação aos processos que influenciam a TSM e as características principais correntes de superfície na região da Macaronésia concluiu-se que a AzC é dominada por fenómenos de mesoescala provocando os movimentos circulares das boias. Também sofre um relaxamento da circulação para Oeste nos meses mais frios e mais influência para Sul nos meses mais quentes. A CaC é a corrente onde se verificou menos variação sazonal e direção constante ao longo do ano. A NEC sofreu uma alteração de latitude para norte nos meses mais quentes e a PC provoca deslocamentos rápidos nas boias. Uma das conclusões mais importantes que surgiu desta questão derivante foi precisamente as áreas onde se verificou maior importância no lançamento das boias para obtenção de dados TSM. E estas são ao largo de Portugal Continental, Arquipélago da Madeira, costa noroeste de África e Arquipélago das Canárias devido ao pouco tempo de permanência de boias nestas regiões.

Em relação aos métodos de obtenção de dados de TSM, os três métodos em estudo mostraram-se diferentes na forma de medição de temperaturas. As boias derivantes medem-na *in-situ* através de termístores. O produto de imagens satélite combina vários resultados de imagens microondas e infravermelhos. Estas medições

são efetuadas à distância através da radiação emitida pela superfície do mar e que é interceptada por sensores presentes nos satélites. Já o modelo numérico, é um modelo que assimila dados de boias fundeadas, derivantes, navios e imagens satélite. Deste modo, acaba por ser já uma combinação dos outros dois métodos. Derivado destas formas de obtenção dos dados e da correspondente profundidade da camada de superfície em que medem a TSM, é espetável que hajam diferenças entre  $TSM_{GDP}$ ,  $TSM_{MOD}$  e  $TSM_{SAT}$ . No entanto, estas verificam-se maiores perante fenómenos de mesoescala, nomeadamente *eddies*, e estes encontram-se mais frequentes nas regiões do Arquipélago dos Açores e Cabo Verde nos meses mais frios.

Num panorama ideal a TSM deveria ser sempre registada *in-situ*. As boias derivantes apenas têm uma cobertura de 77% e como consequência, a utilização única deste método não é suficiente. Estas têm um tempo médio de vida de 13 meses e deste modo, todos os anos tenham de ser lançadas aproximadamente 1000 boias para manter a matriz objetivo do programa (custo associado). Para além disso, uma vez lançadas não podem ser calibradas e estão condicionadas a condições de mar muito adverso. O maior aspeto positivo é medirem os dados *in-situ* que até hoje é a forma de medição que traduz maior confiança e fornece dados em tempo real. Também se verificou que os registos do GDP, uma vez que são medidos no local onde estas se encontram, oferecem pouca quantidade de dados à escala global surgindo a necessidade de recorrer a métodos de interpolação com os erros intrínsecos a este. Por outro lado, verificou-se que 1 das 114 boias possuía uma má calibração durante 349 dias sucessivos, o que manifesta a necessidade de uma utilização conjunta com outros métodos pois mesmo o *ground truth* do estudo teve, embora em baixa percentagem, um erro associado.

Já o MOD necessita de assimilar dados provenientes de satélites e dados *in-situ* de boias e navios, e nunca poderá ser independente deste outros métodos. No entanto, o aspeto mais positivo é que fornece dados instantâneos a todas as horas do dia, e efetua previsões e análises aos seus registos por forma a mitigar os erros. Para além disso, não tem custos associados (apenas os custos dos modelos em que se apoia) e tem uma abrangência global espacial e temporal.

Por último, os dados SAT, este produto de nível 4 que possui sensores de radiação de infravermelhos e micro-ondas obteve no geral bons resultados em condições mais adversas como nebulosidade. No entanto, verificou-se que a qualidade e precisão dos seus registos era superior nos meses mais quentes, de maio a agosto, conforme demonstra a Figura 20. Estes carecem de calibração e validação por dados

---

*in-situ*. Neste estudo, verificou-se que em alguns momentos, a atenuação das variações bruscas de temperatura tendo alguma dificuldade em detetar os picos de subida e descida de temperatura. Este foi aquele onde se verificou valores mais oscilantes, menos precisos e carecem de calibração por dados *in-situ* e têm um alto custo de construção e manutenção.

No presente estudo apurou-se que, para a região da Macaronésia, de um modo geral, as  $TSM_{MOD}$  são mais exatas, isto é, menos diferença entre o valor medido e o valor observado pelo GDP, do que as  $TSM_{SAT}$ . Regionalmente, verificou-se que a este das ilhas canárias, entre 27° e 33°N e 8° e 14°W as  $TSM_{SAT}$  são mais exatas do que as  $TSM_{MOD}$ . Entre 33° e 40°N e 8 e 14°W, e 27° e 33°N e 14° e 26°W os dois métodos obtiveram resultados semelhantes. Em toda a restante área da macaronésia, a exatidão das  $TSM_{MOD}$  é superior às  $TSM_{SAT}$ . No geral, as áreas com piores previsões são entre as latitudes 27° e 33°N e longitudes 26° e 32°W, latitudes 14° e 21° e longitudes 14° e 26°W, e latitudes 21° e 27°N e longitudes 14° e 20°W.

Temporalmente, verificou-se mais exatidão dos dados MOD nos meses de setembro, janeiro, fevereiro, março e junho enquanto nos SAT a exatidão nos meses de janeiro, fevereiro, março, abril e junho.

As boias lançadas pelo NRP Sagres que obtiveram piores resultados foram a 6202656 com piores exatidões MOD (derivou entre latitudes 27° e 40°N e longitudes 8° e 14°W) e a 6202657 com piores exatidões SAT (derivou entre latitudes 14° e 21°N e longitudes 20° e 32°W), o que vai de encontro ao resultado obtido pelas boias todas ao longo de todo o período de estudo.

Em suma, dependendo da aplicação que se queira dar aos registos, cada um dos métodos têm uma elevada eficiência. Se a precisão for a prioridade, a melhor escolha recai na  $TSM_{GDP}$ , caso seja um local com bastante fluxo de boias. No caso de se atribuir maior importância à cobertura geral espacial e temporal, a melhor escolha é as  $TSM_{MOD}$ . Por último, a escolha ideal para ter uma cobertura geral e mais independente é as  $TSM_{SAT}$ .

Como recomendações futuras, deverá ser feita uma comparação entre metodologias completamente independentes, isto é, que não sejam assimiladas (como acontece neste estudo entre o MOD e SAT), e uma comparação entre os erros nas horas diurnas e noturnas, de modo a estudar a variabilidade diurna.





# Bibliografia

- Abdi, H. (2007). The Kendall rank correlation coefficient. *Encyclopedia of Measurement and Statistics*. Sage, Thousand Oaks, CA, 508–510.
- Alvera-Azcárate et al. (2011). Comparison between satellite and in situ sea surface temperature data in the Western Mediterranean Sea. *Ocean Dynamics*, 61(6), 767–778. <https://doi.org/10.1007/s10236-011-0403-x>
- Anaconda. (s.d.). *Anaconda documentation*. <https://docs.anaconda.com/anaconda/> acedido em março de 2021
- Barton, E. et al. (1998). The transition zone of the Canary Current upwelling region. *Progress in Oceanography*, 41(4), 455–504. [https://doi.org/10.1016/S0079-6611\(98\)00023-8](https://doi.org/10.1016/S0079-6611(98)00023-8)
- Barton, E. (2001). Canary and Portugal currents. *Ocean Currents*, 330–339. <https://doi.org/10.1016/B978-0-12-813081-0.00360-8>
- Brennan, M. o. (2004). Ultrashort optical pulse generation with a mode-locked long-wavelength (1075–1085 nm) InGaAs-GaAs semiconductor laser. *IEEE Photonics Technology Letters*, 16(8), 1798–1800. <https://doi.org/10.1109/LPT.2004.831320>
- Brügge, B. (1995). Near-surface mean circulation and kinetic energy in the central North Atlantic from drifter data. *Journal of Geophysical Research: Oceans*, 100(C10), 20543–20554. <https://doi.org/10.1029/95JC01501>
- Chai, T. & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific Model Development*, 7(1), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chaves, F. (2015). *Estudo climatológico da nortada na costa oeste da Península Ibérica* (tese de doutoramento).
- CMEMS. (s.d.). *Global ocean 1/4° physics analysis and forecast updated daily*. [https://resources.marine.copernicus.eu/?option=com\\_csw&view=details&product\\_id=GLOBAL\\_ANALYSISFORECAST\\_PHY\\_CPL\\_001\\_015](https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=GLOBAL_ANALYSISFORECAST_PHY_CPL_001_015) (acedido em março de 2021)
- Donlon, C. et al. (2009). Successes and challenges for the modern sea surface temperature observing system. *Community White Paper for OceanObs*, 9, 249–258. <https://doi.org/10.5270/oceanobs09.cwp.24>

- Donlon, C. et al. (2001). The character of skin and subsurface sea surface temperature. *WORLD METEOROLOGICAL ORGANIZATION-PUBLICATIONS-WMO TD*, 298–302.
- Donlon, C. et al. (2007). The global ocean data assimilation experiment high-resolution sea surface temperature pilot project. *Bulletin of the American Meteorological Society*, 88(8), 1197–1214.
- Emery, W. et al. (2001). Estimating sea surface temperature from infrared satellite and in situ temperature data. *Bulletin of the American Meteorological Society*, 82(12), 2773–2786.
- Fernandes, M. et al. (2005). Oceanographic characterization of the Cape Verde region using multisensor data. *Envisat & ERS Symposium*, 572.
- Figueiredo, D. et al. (2009). Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r). *Revista Política Hoje*, 18(1), 115–146.
- Gentemann, C. L. et al. (2009). Accuracy of satellite sea surface temperatures at 7 and 11 GHz. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3), 1009–1018.
- GHRSSST. (s.d.). *GHRSSST – The Group for High Resolution Sea Surface Temperature*. <https://www.ghrsst.org/ghrsst-data-services/products/> acessado em março de 2021
- Gibelin, A. & Déqué, M. (2003). Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model. *Clim. Dyn.*, 20, 327–. <https://doi.org/10.1007/s00382-002-0277-1>
- Gill, A. E. (2016). *Atmosphere—ocean dynamics*. Elsevier.
- Giorgi, F. (2006). Climate change hot-spots. *Geophysical Research Letters*, 33(8). <https://doi.org/10.1029/2006GL025734>
- Guan, L. & Kawamura, H. (2004). Merging satellite infrared and microwave SSTs: Methodology and evaluation of the new SST. *Journal of Oceanography*, 60(5), 905–912.
- Guiavarc’h, C. et al. (2019). Assessment of ocean analysis and forecast from an atmosphere–ocean coupled data assimilation operational system. *Ocean Science*, 15(5), 1307–1326. <https://doi.org/10.5194/os-15-1307-2019>
- IH. (s.d.). *Projeto Científico entre o Instituto Hidrográfico, o NRP Sagres e a NOAA*. <https://www.hidrografico.pt/noticia/614> acessado em maio de 2020
- Jia, Y. (2000). Formation of an Azores Current due to Mediterranean overflow in a modeling study of the North Atlantic. *Journal of Physical Oceanography*, 30(9), 2342–2358. [https://doi.org/10.1175/1520-0485\(2000\)030<2342:FOAACD>2.0.CO;2](https://doi.org/10.1175/1520-0485(2000)030<2342:FOAACD>2.0.CO;2)

- Kleissl, J. (2013). *Solar energy forecasting and resource assessment*. Academic Press.
- Lauvset, S. & Tanhua, T. (2015). A toolbox for secondary quality control on ocean chemistry and hydrographic data. *Limnology and Oceanography: Methods*, 13(11), 601–608.
- Le Menn, M. et al. (2019). Development of surface drifting buoys for fiducial reference measurements of sea-surface temperature. *Frontiers in Marine Science*, 6, 578.
- Lea, D. et al. (2015). Assessing a new coupled data assimilation system based on the Met Office coupled atmosphere–land–ocean–sea ice model. *Monthly Weather Review*, 143(11), 4678–4694. <https://doi.org/10.1175/MWR-D-15-0174.1>
- Machín, F. et al. (2006). Mass fluxes in the Canary Basin. *Progress in Oceanography*, 70(2-4), 416–447.
- Mann, C. (1967). The termination of the Gulf Stream and the beginning of the North Atlantic Current. *Deep Sea Research and Oceanographic Abstracts*, 14(3), 337–359.
- Marshall, J. & Schott, F. (1999). Open-ocean convection: Observations, theory, and models. *Reviews of geophysics*, 37(1), 1–64.
- Martin, M. et al. (2012). Group for High Resolution Sea Surface temperature (GHR SST) analysis fields inter-comparisons. Part 1: A GHR SST multi-product ensemble (GMPE). *Deep Sea Research Part II: Topical Studies in Oceanography*, 77, 21–30.
- Mason, E. et al. (2005). An overview of the literature concerning the oceanography of the eastern North Atlantic region. *Relatórios Científicos e Técnicos IPIMAR Serie Digital*, 33, 59.
- Matthews, J. (2013). Comparing historical and modern methods of sea surface temperature measurement—Part 1: Review of methods, field comparisons and dataset adjustments. *Ocean Science*, 9(4), 683–694.
- Motyzhev, S. et al. (2012). Increasing of effectiveness and reliability of data from drifting buoys. *DBCP Technical Document*, (45).
- NOAA. (s.d.-a). *Global Drifter Program drifting buoy collection*. [http://osmc.noaa.gov/erddap/tabledap/gdp\\_interpolated\\_drifter.subset](http://osmc.noaa.gov/erddap/tabledap/gdp_interpolated_drifter.subset)
- NOAA. (s.d.-b). NOAA. [https://www.aoml.noaa.gov/phod/gdp/images/GDP\\_Brochure\\_Digital\\_v2.pdf](https://www.aoml.noaa.gov/phod/gdp/images/GDP_Brochure_Digital_v2.pdf) acedido em março de 2021
- NOAA. (s.d.-c). *Physical Oceanography Division(PhOD) - Global Drifter Program*. <https://www.aoml.noaa.gov/phod/gdp/> acedido em março de 2021

- Peña-Izquierdo et al. (2012). The continental slope current system between Cape Verde and the Canary Islands. *Sci. Mar*, 76, 65–78. <https://doi.org/10.3989/scimar.03607.18C>
- Pérez-Hernández, M. et al. (2013). The source of the Canary current in fall 2009. *Journal of Geophysical Research: Oceans*, 118(6), 2874–2891.
- Relvas, P. et al. (2007). Physical oceanography of the western Iberia ecosystem: latest views and challenges. *Progress in Oceanography*, 74(2-3), 149–173.
- Rodrigues, M. (2011). O tratamento e análise de dados. *Metodologia para a investigação social*, 179–230.
- Sælen, O. H. & Aas, E. (2012). *Lecture notes in physical oceanography*.
- Stewart, R. H. (2008). *Introduction to physical oceanography*. Robert H. Stewart.
- Teles-Machado, A. et al. (2007). On the onset of the Gulf of Cadiz Coastal Countercurrent. *Geophysical Research Letters*, 34(12). <https://doi.org/10.1029/2007GL030091>
- Volkov, D. L. & Fu, L.-L. (2010). On the reasons for the formation and variability of the Azores Current. *Journal of Physical Oceanography*, 40(10), 2197–2220. <https://doi.org/10.1175/2010JPO4326.1>
- Wilk, M. & Gnanadesikan, R. (1968). Probability plotting methods for the analysis for the analysis of data. *Biometrika*, 55(1), 1–17. <https://doi.org/10.1093/biomet/55.1.1>
- Yug, M. (s.d.). *Marlin Yug*. <https://marlin-yug.com/productfile/SVP-B.pdf> acedido em março de 2021
- Zhou, M. et al. (2000). Surface currents in the Canary Basin from drifter observations. *Journal of Geophysical Research: Oceans*, 105(C9), 21893–21911. <https://doi.org/10.1029/2000jc900096>

## Apêndice A - Processo de obtenção do ficheiro dos dados sujeitos a análise

Os registos dos três conjuntos de dados foram removidos *online* e denominam-se da seguinte forma:

- MetO-GLO-PHY-CPL-hi-TEM\_1623052658711.nc, que contém os dados do modelo numérico;
- METOFFICE-GLO-SST-L4-NRT-OBS-GMPE-V3\_1623062334481.nc, que contém os dados do produto dos satélites;
- gdp\_interpolated\_drifter\_c65f\_788d\_de52.csv, que contém os dados das boias derivantes.

MOD		SAT		GDP	
Index	0	Index	0	Index	time
0	2019-06-29 23:40:00	0	2019-06-30 12:00:00	1	2019-06-30 00:00:00
1	2019-06-30 00:40:00	1	2019-07-01 12:00:00	2	2019-06-30 06:00:00
2	2019-06-30 01:40:00	2	2019-07-02 12:00:00	3	2019-06-30 12:00:00
3	2019-06-30 02:40:00	3	2019-07-03 12:00:00	4	2019-06-30 18:00:00
4	2019-06-30 03:40:00	4	2019-07-04 12:00:00	5	2019-07-01 00:00:00
5	2019-06-30 04:40:00	5	2019-07-05 12:00:00	6	2019-07-01 06:00:00
6	2019-06-30 05:40:00	6	2019-07-06 12:00:00	7	2019-07-01 12:00:00
7	2019-06-30 06:40:00	7	2019-07-07 12:00:00	8	2019-07-01 18:00:00
8	2019-06-30 07:40:00	8	2019-07-08 12:00:00	9	2019-07-02 00:00:00
9	2019-06-30 08:40:00	9	2019-07-09 12:00:00	10	2019-07-02 06:00:00
10	2019-06-30 09:40:00	10	2019-07-10 12:00:00	11	2019-07-02 12:00:00
11	2019-06-30 10:40:00	11	2019-07-11 12:00:00	12	2019-07-02 18:00:00
12	2019-06-30 11:40:00	12	2019-07-12 12:00:00	13	2019-07-03 00:00:00
13	2019-06-30 12:40:00	13	2019-07-13 12:00:00	14	2019-07-03 06:00:00
14	2019-06-30 13:40:00	14	2019-07-14 12:00:00	15	2019-07-03 12:00:00
15	2019-06-30 14:40:00	15	2019-07-15 12:00:00	16	2019-07-03 18:00:00

FIGURA A.1: Serie temporal dos três conjuntos de dados em que a primeira coluna corresponde aos dados MOD, a segunda aos dados SAT, e a terceira aos GDP.

Por forma a efetuar o corte dos dois hipercubos (MOD e SAT) por posição geográfica das boias GDP, criou-se um código em python que coloca-se os três conjuntos de dados referentes à mesma posição geográfica, uma vez que têm séries temporais diferentes, de acordo com a figura A.1.

## A.1 Código de criação do Excel

Através da função *datetime* e *timedelta*, foi possível retirar do universo de dados que foram removidos *online*, apenas as horas que se intersetavam nos três conjuntos de dados. Como os dados MOD são apresentados ao minuto 40, arredondou-se por forma a serem intersetáveis com os outros dois conjuntos de dados. Após várias ações de verificar se os arranjos tinham o mesmo tamanho, criou-se a tabela *cdata*, a partir da tabela proveniente do *Excel* onde se contra os dados GDP, que vai auxiliar na compilação das variáveis.

```

1 from netCDF4 import Dataset, num2date
2 import numpy as np
3
4 file1='MetO-GLO-PHY-CPL-hi-TEM_1623052658711.nc'
5 file2='METOFFICE-GLO-SST-L4-NRT-OBS-GMPE-V3_1623062334481.nc'
6 file3='gdp_interpolated_drifter_c65f_788d_de52.csv'
7
8 data1_sst=Dataset(file1, mode='r')
9 data2_sst=Dataset(file2, mode='r')
10
11 import pandas as pd
12 data3=pd.read_csv(file3)[1:]
13
14 time_sst_1=data1_sst.variables['time']
15 time_sst_2=data2_sst.variables['time']
16
17 time_mod_sst_1 = num2date (time_sst_1,units='seconds since 2017-07-11
    00:00',calendar='gregorian')
18 time_mod_sst_1 = pd.Series(time_mod_sst_1)
19
20 time_mod_sst_2 = num2date (time_sst_2,units='seconds since 1981-01-01
    00:00',calendar='gregorian')
21 time_mod_sst_2 = pd.Series(time_mod_sst_2)
22
23 time_data3 = data3.time
24 time_mod_data3 = time_data3.astype('datetime64[s]')
25
26 from datetime import datetime, timedelta
27 # make all datetime to better call it
28 date1 = []
29
30 # Atraves do timedelta, arredonda-se 0040 para 0100 e assim
    sucessivamente

```

```

31
32 for i in range(len(time_mod_sst_1)):
33     a = time_mod_sst_1[i]
34     reg = datetime(a.year,a.month,a.day,a.hour)+timedelta(hours=1)#
    round up
35     date1.append(reg)
36
37 date2 = []
38
39 for i in range(len(time_mod_sst_2)):
40     b = time_mod_sst_2[i]
41     reg = datetime(b.year,b.month,b.day,b.hour)
42     date2.append(reg)
43
44 date3 = []
45
46 for i in range(1,len(time_mod_data3)+1):
47     c = time_mod_data3[i]
48     reg = datetime(c.year,c.month,c.day,c.hour)
49     date3.append(reg)
50
51 # Para confirmar
52 print(len(date3)==len(time_mod_data3))
53 print(len(date2)==len(time_mod_sst_2))
54 print(len(date1)==len(time_mod_sst_1))
55 data3['datetime'] = np.array(date3) # Turn it into an array
56 date1 = np.array(date1)
57 date2 = np.array(date2)
58 date3 = np.array(date3)
59
60 # A partir do data3 cria-se o cdata (concatenated data)
61 cdata = data3.copy().drop(columns='time')
62 cdata['longitude']=cdata['longitude'].astype(float)
63 cdata['latitude']=cdata['latitude'].astype(float)
64 cdata['temp']=cdata['temp'].astype(float) # de string para float

```

Para se obter o registo mais próximo de cada registo de boias, aplicou-se o método do vizinho mais próximo que utiliza distâncias euclidianas.

```

1 # funcao para obter o registo mais proximo de cada registo de boias
2
3 local = [cdata.latitude[1],cdata.longitude[1]]
4 lat_2 = data2_sst.variables['lat'][:].data

```

```

5 lon_2 = data2_sst.variables['lon'][:].data
6 coordinates_data2 = np.array(np.meshgrid(lat_2,lon_2)).T.reshape(-1,2)
7 from nearest_local import nearest_local as nl
8 print(nl(local,coordinates_data2))
9
10 #Como o SAT apenas tem registos as 1200h, apenas permaneceram estes
    registos
11
12 hour = []
13 for i in range(1, len(cdata)+1):
14     hour.append(cdata.datetime[i].hour)
15 cdata['hour']=hour
16 cdata_12 = cdata[cdata.hour==12]
17
18 i=0
19 lat_1 = data1_sst.variables['lat']
20 lat_2 = data2_sst.variables['lat']
21 lon_1 = np.array(data1_sst.variables['lon'])-360
22 lon_2 = data2_sst.variables['lon']
23 coordinates_data1 = np.array(np.meshgrid(lat_1,lon_1)).T.reshape(-1,2)
24 coordinates_data2 = np.array(np.meshgrid(lat_2,lon_2)).T.reshape(-1,2)
25 lat_data1 = []
26 lat_data2 = []
27 lon_data1 = []
28 lon_data2 = []
29 sst_1 = []
30 sst_2 = []
31
32 # cicle
33 for i in range(len(cdata_12)):
34     line = cdata_12.iloc[i] #call the line of data3
35     time_selected = line.datetime # draw date of that line
36     lat_selected = line.latitude # draw lat
37     lon_selected = line.longitude # draw lon
38     local = [lat_selected,lon_selected] # defines the local
39     ind_t1 = np.where(date1==time_selected)[0][0] # find index of date
        in data1
40     ind_t2 = np.where(date2==time_selected)[0][0] # find index of data
        in data2
41     near1 = nl(local,coordinates_data1) # index and nearest local from
        data1
42     near2 = nl(local,coordinates_data2) # index and nearest local from
        data2
43     lat_n1 = near1[1][0]
44     ind_lat1 = np.where(lat_1==lat_n1)[0][0]

```



```

45     lat_n2 = near2[1][0]
46     ind_lat2 = np.where(lat_2==lat_n2)[0][0]
47     lon_n1 = near1[1][1]
48     ind_lon1 = np.where(lon_1==lon_n1)[0][0]
49     lon_n2 = near2[1][1]
50     ind_lon2 = np.where(lon_2==lon_n2)[0][0] # all indexes for lat and
lon in both data
51     lat_data1.append(lat_n1)
52     lat_data2.append(lat_n2)
53     lon_data1.append(lon_n1)
54     lon_data2.append(lon_n2)
55
56     sst_2.append(data2_sst.variables['analysed_sst'][ind_t2,ind_lat2,
ind_lon2]) # saves sst in Kelvin's
57     sst_1.append(data1_sst.variables['thetao'][ind_t1,ind_lat1,ind_lon1])
58     if i%100==0: print(i)
59     len(cdata_12)
60     cdata_12['lat1']=lat_data1.copy()
61     cdata_12['lon1']=lon_data1.copy()
62     cdata_12['sst_1']=np.array(sst_1.copy())
63     cdata_12['lat2']=lat_data2.copy()
64     cdata_12['lon2']=lon_data2.copy()
65     cdata_12['sst_2']=np.array(sst_2.copy())-273.15
66     cdata_12.to_excel('data_merge.xlsx')

```

Para cada localização, o código do método do vizinho mais próximo criado, calcula a distância euclidiana das coordenadas dessa localização e retorna a distância mínima.

```

1 import numpy as np
2
3 def nearest_local(local, coordinates_data):
4     distance_to_local = np.sqrt(np.square(coordinates_data-local).sum(
axis=1))
5     ind_local = np.argmin(distance_to_local)
6     return ind_local, coordinates_data[ind_local]

```

Após a aplicação do código supramencionado, obteve-se o ficheiro Excel data\_merge.xlsx representado na figura A.2.

	WMO	longitude	latitude	temp	datetime	lat1	lon1	sst_1	lat2	lon2	sst_2
3	1300613	-21,657	28,095	23,184	2019-06-30 12:00:00	28	-21,75	23,194	28,125	-21,625	23,26
7	1300613	-21,659	28,15	23,243	2019-07-01 12:00:00	28,25	-21,75	23,169	28,125	-21,625	23,17001
11	1300613	-21,678	28,207	23,201	2019-07-02 12:00:00	28,25	-21,75	23,073	28,125	-21,625	23,1
15	1300613	-21,736	28,273	23,098	2019-07-03 12:00:00	28,25	-21,75	22,924	28,375	-21,625	22,95001
19	1300613	-21,793	28,338	23,018	2019-07-04 12:00:00	28,25	-21,75	23,079	28,375	-21,875	22,92999
23	1300613	-21,827	28,39	23,045	2019-07-05 12:00:00	28,5	-21,75	23,187	28,375	-21,875	23,1
27	1300613	-21,817	28,426	23,225	2019-07-06 12:00:00	28,5	-21,75	23,111	28,375	-21,875	23,14999
31	1300613	-21,847	28,471	23,233	2019-07-07 12:00:00	28,5	-21,75	23,219	28,375	-21,875	23,20001
35	1300613	-21,888	28,507	23,268	2019-07-08 12:00:00	28,5	-22	23,432	28,625	-21,875	23,11999
39	1300613	-21,924	28,554	23,282	2019-07-09 12:00:00	28,5	-22	23,262	28,625	-21,875	23,1
43	1300613	-21,965	28,583	23,245	2019-07-10 12:00:00	28,5	-22	23,274	28,625	-21,875	23,13
47	1300613	-22,014	28,615	23,369	2019-07-11 12:00:00	28,5	-22	23,409	28,625	-22,125	23,27999
51	1300613	-22,098	28,652	23,337	2019-07-12 12:00:00	28,75	-22	23,254	28,625	-22,125	23,39999
55	1300613	-22,176	28,702	23,387	2019-07-13 12:00:00	28,75	-22,25	23,345	28,625	-22,125	23,39999
59	1300613	-22,268	28,768	23,412	2019-07-14 12:00:00	28,75	-22,25	23,389	28,875	-22,375	23,36999
63	1300613	-22,369	28,845	23,389	2019-07-15 12:00:00	28,75	-22,25	23,164	28,875	-22,375	23,26
67	1300613	-22,476	28,905	23,38	2019-07-16 12:00:00	29	-22,5	23,223	28,875	-22,375	23,26
71	1300613	-22,604	28,97	23,314	2019-07-17 12:00:00	29	-22,5	23,022	28,875	-22,625	23,13
75	1300613	-22,724	29,041	23,262	2019-07-18 12:00:00	29	-22,75	23,148	29,125	-22,625	23,07998
79	1300613	-22,822	29,122	23,22	2019-07-19 12:00:00	29	-22,75	23,287	29,125	-22,875	23,20999
83	1300613	-22,913	29,186	23,261	2019-07-20 12:00:00	29,25	-23	23,31	29,125	-22,875	23,22

FIGURA A.2: Imagem retirada do ficheiro Excel data\_merge.xlsx

## Apêndice B - Descrição do tratamento dos dados

Após a obtenção do ficheiro `data_merge.xlsx`, verificou-se que 0,5% dos registos tinham campos em branco pelo que se eliminou esses registos ficando com 19119 dados para analisar.

Uma vez que se trata de uma quantidade de amostragem elevada, aplicou-se o teste de Kolmogorov-Smirnov e estudou-se a dispersão dos resíduos dos registos por forma a estudar problemas com as boias e, como se veio a verificar, casos anómalos que estavam a piorar os resultados.

Um dos critérios para a aplicação do teste de Kolmogorov-Smirnov é que a amostra seja superior a 30 dados, o que se verifica. Este teste permitiu verificar a qualidade do ajuste entre  $TSM_{GDP}$  e as  $TSM_{MOD}$  e  $TSM_{SAT}$ , e consequentemente a concordância entre os registos dos diferentes métodos.

Os resultados que se obtiveram foi de  $TSM_{MOD}$  foi 0,3298 enquanto a  $TSM_{SAT}$  foi 0,3465, o que significa que, de um modo geral, o MOD se encontra mais ajustado à *ground-truth* do que o SAT. Para além disso, o p-value em ambos é igual a zero pelo que se rejeitou a hipótese de distribuição normal em todos os conjuntos de dados. Esta conclusão foi complementada com o gráfico probabilístico (QQplot) presente na Figura B.1. A linearidade dos pontos sugere que os valores de TSM do três conjuntos de dados não são normais (Wilk e Gnanadesikan, 1968).

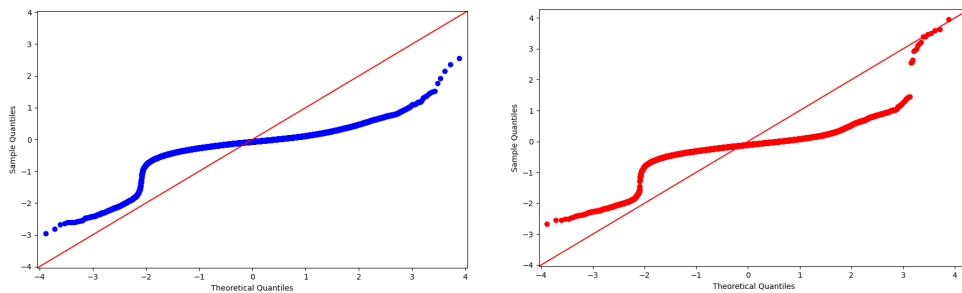


FIGURA B.1: Gráficos probabilísticos de quantis.

Ao estudar a dispersão dos dados de TSM, encontrou-se uma nuvem de pontos afastada dos restantes dados. O declive da reta de regressão dos dados MOD é 0,983 e 0,984 nos dados SAT, conforme representado na Figura B.2. Apesar da nuvem, existe uma elevada concordância nos dois casos com  $R^2$  superiores a 0,98 mas com tendência dos valores tanto MOD como SAT estarem abaixo da reta de regressão.

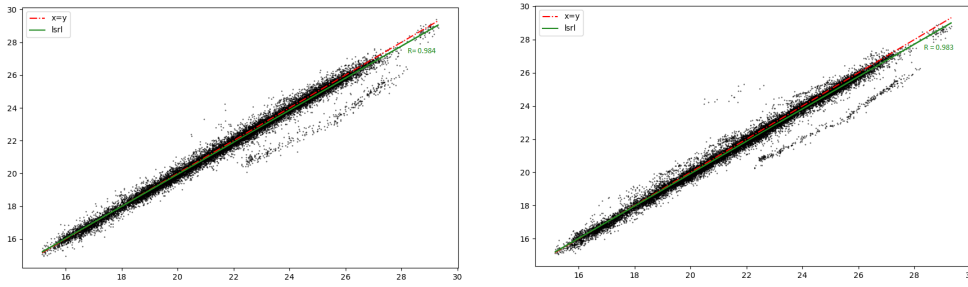


FIGURA B.2: Gráfico de dispersão com as respectivas retas de regressão a verde e retas  $x=y$  a tracejado a vermelho, sendo o gráfico da esquerda relativo aos dados MOD e o da direita aos dados SAT.

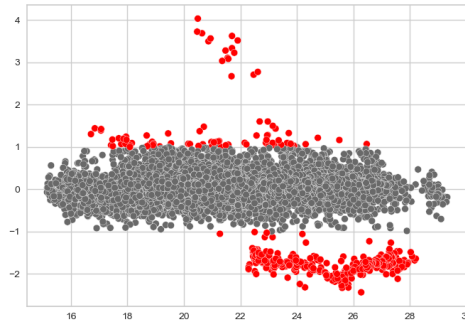
Para melhor se compreender a nuvem de dados, separou-se os dados com diferenças de temperatura superiores a  $1^{\circ}\text{C}$  relativamente àquelas medidas *in-situ* e obteve-se os gráficos representados na Figura B.3. Nesta condição, existem 410 dados MOD e 449 dados SAT, sendo que correspondem maioritariamente (percentagens de erro superiores a 5%) a duas boias: 4101718 (349 vezes) e 4201517 (39 vezes), conforme Figura B.4.

Uma vez identificada as boias que tinham percentagens de erro superiores a 5%, analisou-se as boias por forma a perceber se devem ser removidas ou mantidas no estudo.

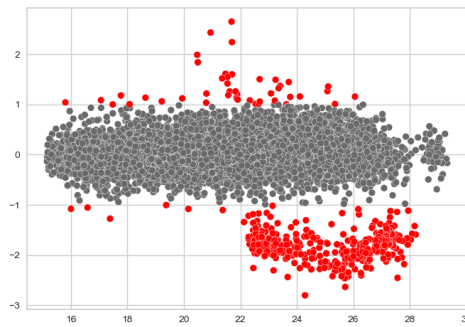
A boia 4101718 possui registos entre 30 de junho de 2019, na posição  $32.7^{\circ}\text{N}$  e  $26.8^{\circ}\text{W}$ , e 12 de junho de 2020, na posição  $24.8^{\circ}\text{N}$  e  $31.9^{\circ}\text{W}$ , e circulou entre as áreas E, F e I, conforme representado na Figura B.5. Dos 349 registos desta boia nenhuma  $TSM_{MOD}$  ou  $TSM_{SAT}$  coincidiu com  $TSM_{GDP}$ , de acordo com a Figura B.6, pelo que se eliminou os registos desta boia.

Relativamente à boia 4201517, o seu trajeto foi nas áreas A e E, conforme a Figura B.7. Ao contrário do caso anterior, os dados SAT foram consistentemente superiores aos restantes. De acordo com a NOAA, esta boia foi lançada no dia 31 de maio de 2017 na posição  $24^{\circ}46.9'\text{N}$  e  $085^{\circ}37.7'\text{W}$ . No entanto, para a área de estudo apenas temos 194 dados entre 22 de outubro de 2019 a 2 de maio de 2020, pelo que se conclui que a boia circulou fora da AI durante os restantes dias. Ao longo do tempo, a tendência da temperatura foi diminuir até ao dia 17 de abril<sup>11</sup>, onde houve uma inversão e as temperaturas aumentaram gradualmente. Embora a diferença máxima entre dados  $TSM_{GDP}$  e  $TSM_{SAT}$  é  $1.4^{\circ}\text{C}$ , os dados  $TSM_{MOD}$

<sup>11</sup>O valor mínimo de  $TSM_{GDP}$  foi atingido no dia 17. No entanto, na  $TSM_{MOD}$  foi no dia 18 e na  $TSM_{SAT}$  foi no dia 16 de abril.



(A) MOD



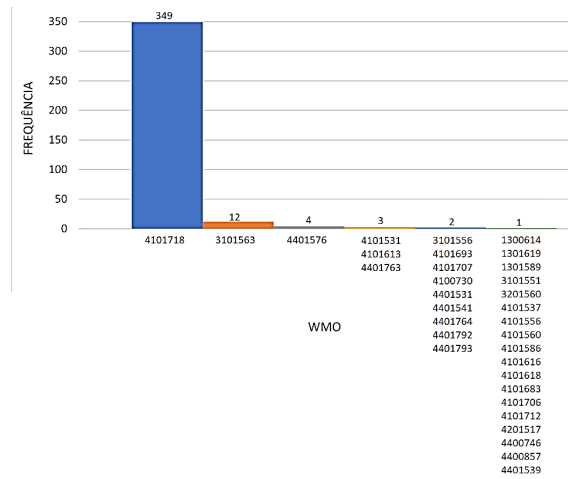
(B) SAT

FIGURA B.3: Gráfico de dispersão dos resíduos MOD (em cima) e SAT (em baixo) com os valores superiores a  $1^{\circ}\text{C}$  a vermelho.

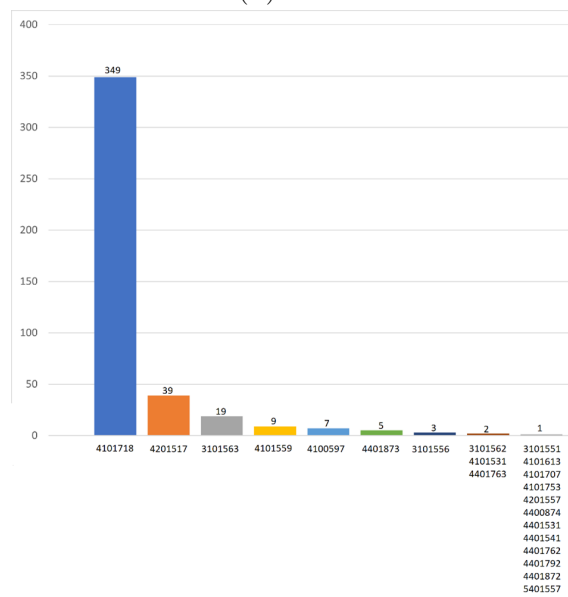
tanto estão mais próximos do primeiro como do segundo. Por esta razão, os dados desta boia não foram eliminados.

Os elementos estatísticos RMSE, MAE, MBE foram calculados para os dados totais antes da eliminação dos registos da boia referida, em específico para as 2 boias supramencionadas e para os dados totais após a eliminação dos registos da boia. Estes permitiram confirmar que a boia 4101718 estava a influenciar negativamente a amostra geral e que tinha valores anormais quando calculados os elementos estatísticos. Após a remoção, verificou-se uma diminuição de erros em todos os elementos estatísticos, de acordo com a figura

Por último, estudou-se nestas duas boias, se havia relação entre a distância entre a localização das boias e a localização das restantes bases de dados e entre a TSM e a consequência de erros e não foi identificada correlação.



(A) MOD



(B) SAT

FIGURA B.4: Histogramas da frequência de diferenças superiores a  $1^{\circ}\text{C}$  entre dados MOD(A) e SAT(B) em relação aos GDP

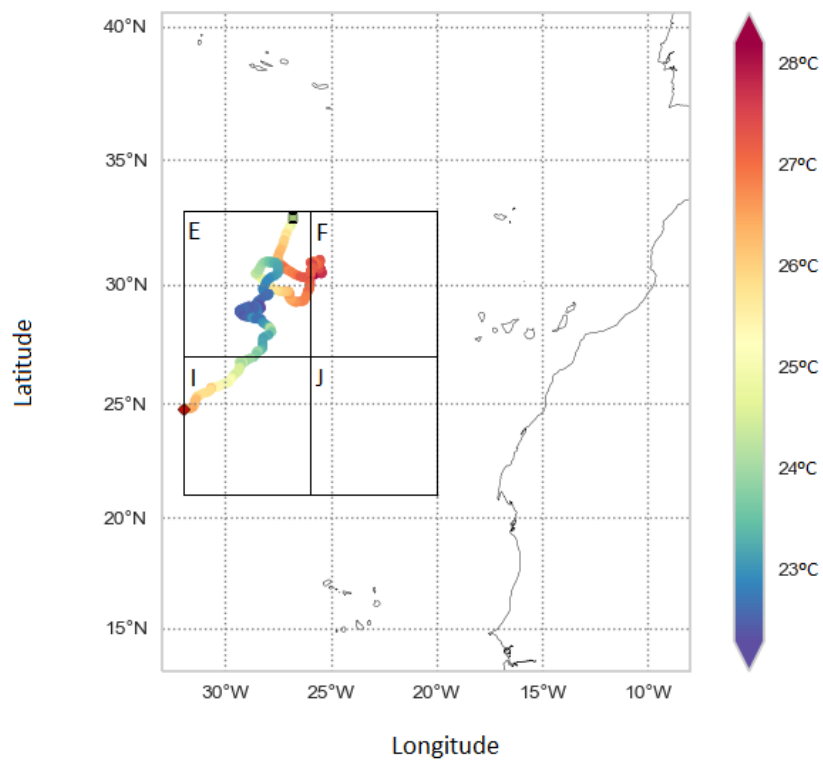


FIGURA B.5: Posições geográficas da boia 4101718, entre 2020. A primeira e última posição estão marcadas por um círculo laranja.

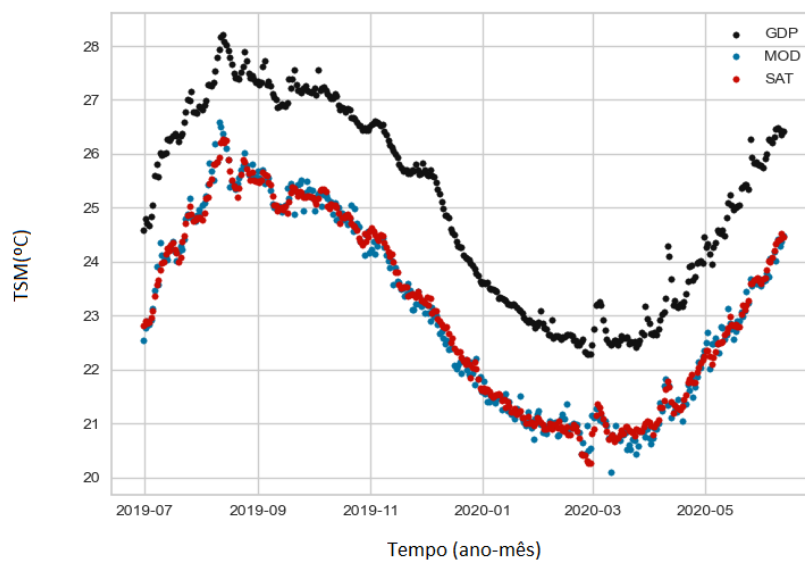


FIGURA B.6: Variação temporal de dados da boia 4101718 de TSM GDP, SAT e MOD.

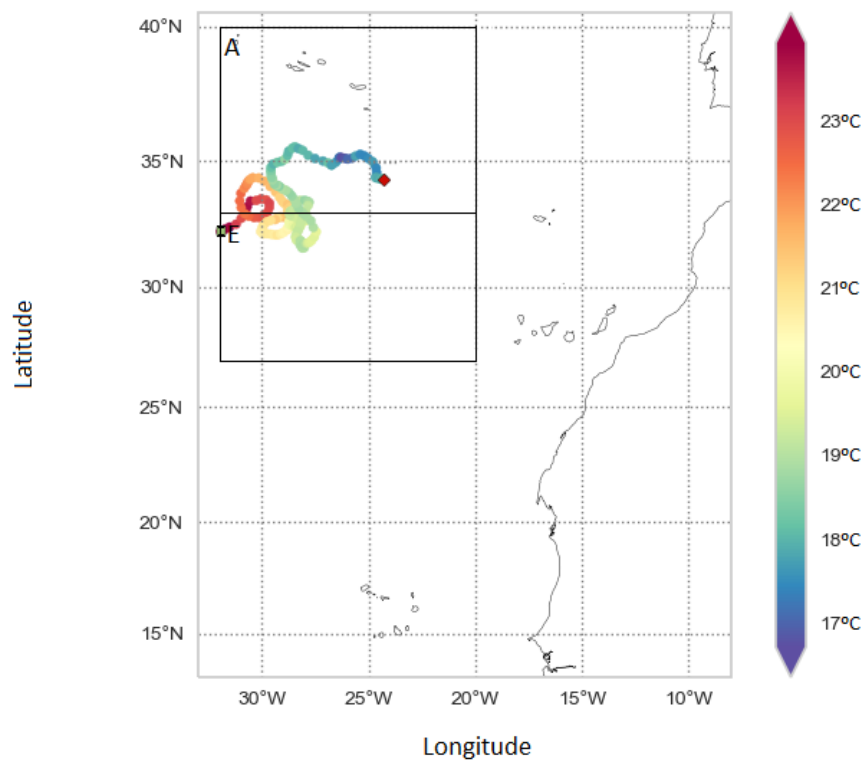


FIGURA B.7: Posições geográficas da boia 4201517, entre 22 de outubro a 2 de maio de 2020. A primeira e última posição estão marcadas por um círculo laranja.

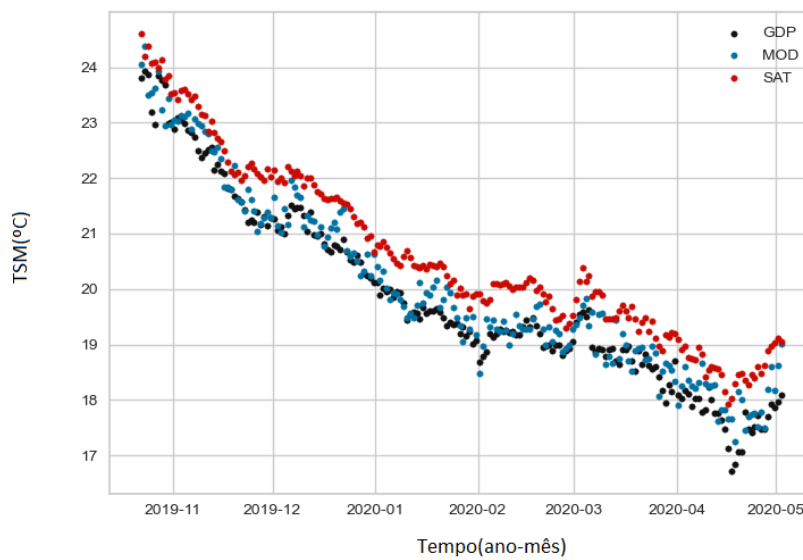


FIGURA B.8: Variação temporal de dados da boia 4201517 de TSM GDP, SAT e MOD .



	Dados totais	Boia 4101718	Boia 4201517	Dados totais refinados
<b>RMSE<sub>MOD</sub></b>	0,363	2.017	0.348	0,245
<b>RMSE<sub>SAT</sub></b>	0,383	1.992	0.794	0,277
<b>MBE<sub>MOD</sub></b>	-0,097	-1,99	0,19	-0,06
<b>MBE<sub>SAT</sub></b>	-0,126	-1.98	0,76	-0,09
<b>MAE<sub>MOD</sub></b>	0.209	1,999	0,274	0,176
<b>MAE<sub>SAT</sub></b>	0.229	1,982	0,761	0,197

TABELA B.1: Tabela resumo dos elementos estatísticos RMSE, MAE, MBE das boias 4201517, 4101718, dos registos globais e dos registos globais após a remoção dos registos da boia 4101718.



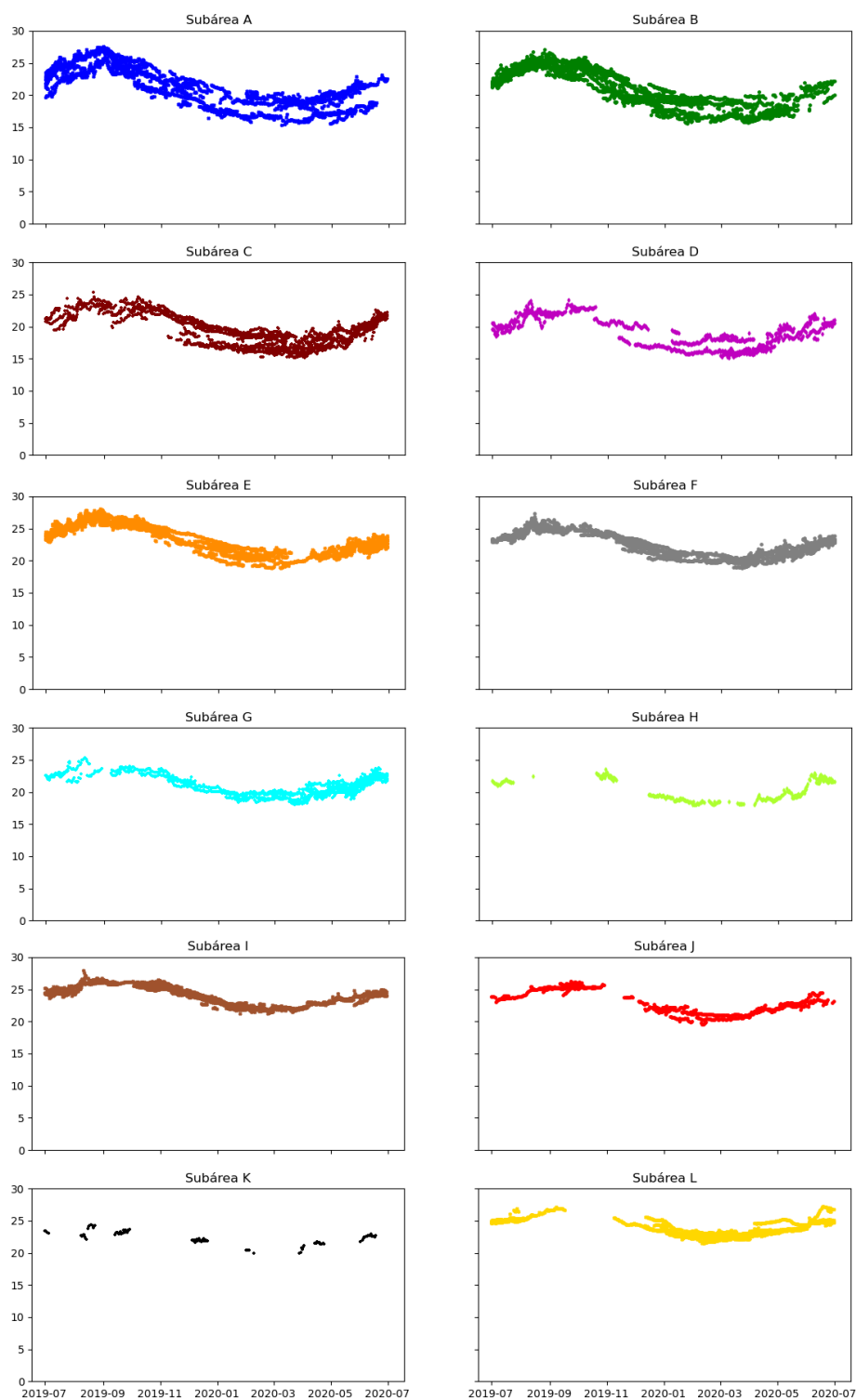
## Apêndice C - Variação temporal dos registos TSM

Conforme representado na Tabela C.1, que se encontra com uma sequência decrescente de número de registos, as subáreas D, G, H, J, K, L, M e N são as que têm menos de 1500 registos ao longo de um ano. A partir desta, também é possível apurar que a área B tem o maior número de registos, mas não é aquela com maior número de boias com registos. Também as áreas H e N possuem o mesmo número de boias com registos, mas a área H possui mais 193 registos do que a outra. Assim conclui-se que uma área que apresente um maior número de boias não tem necessariamente de ter maior número de registos.

Subáreas	Número de registos	Número de boias
<b>B</b>	2778	40
<b>E</b>	2497	38
<b>A</b>	2312	43
<b>C</b>	2271	31
<b>F</b>	2071	32
<b>I</b>	1502	23
<b>G</b>	1470	29
<b>L</b>	1314	24
<b>D</b>	850	14
<b>J</b>	793	20
<b>M</b>	711	18
<b>H</b>	292	10
<b>N</b>	162	5
<b>K</b>	99	10

TABELA C.1: Número de registos e número de boias por subárea geográfica, para o período entre 1 de julho 2019 e 30 de junho 2020.

Relativamente à distribuição das TSM por subárea, as que possuem registos com médias de TSM superiores a 23°C são as E, I, J, L e N enquanto as que registam médias inferiores a 20°C são as C e D (Figura C.1).



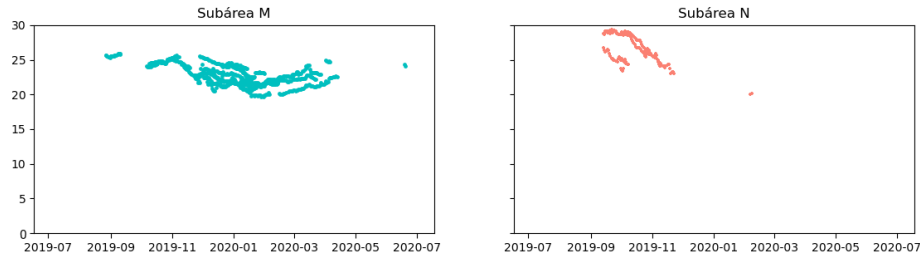


FIGURA C.1: Variação temporal dos registos  $TSM_{GDP}$  por cada uma das subáreas definidas no Capítulo 1.

As subáreas A, B e C têm as amplitudes mais elevadas sendo que são superiores a  $10^{\circ}\text{C}$ . Também se verifica que as subáreas H, J, K, L, M não têm registos contínuos durante todo o período de estudo. Temporalmente, verificou-se mais exatidão dos dados MOD nos meses de setembro, janeiro, fevereiro, março e junho enquanto nos SAT mais exatidão nos meses de janeiro, fevereiro, março, abril e junho (Figura C.2 e figura C.3).

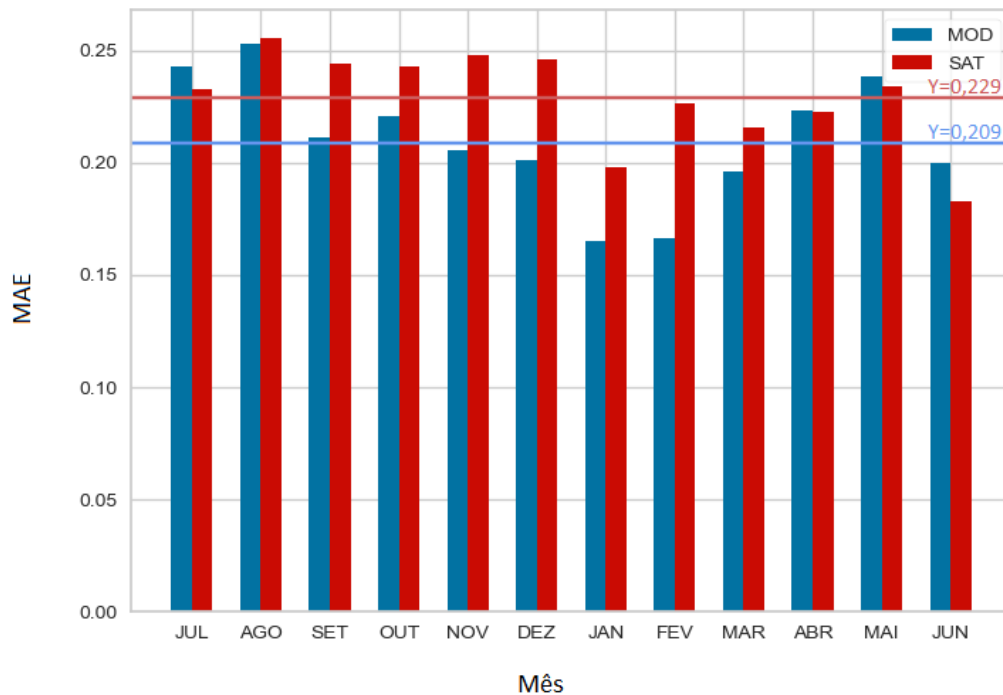


FIGURA C.2: Variação do MAE por mês de estudo

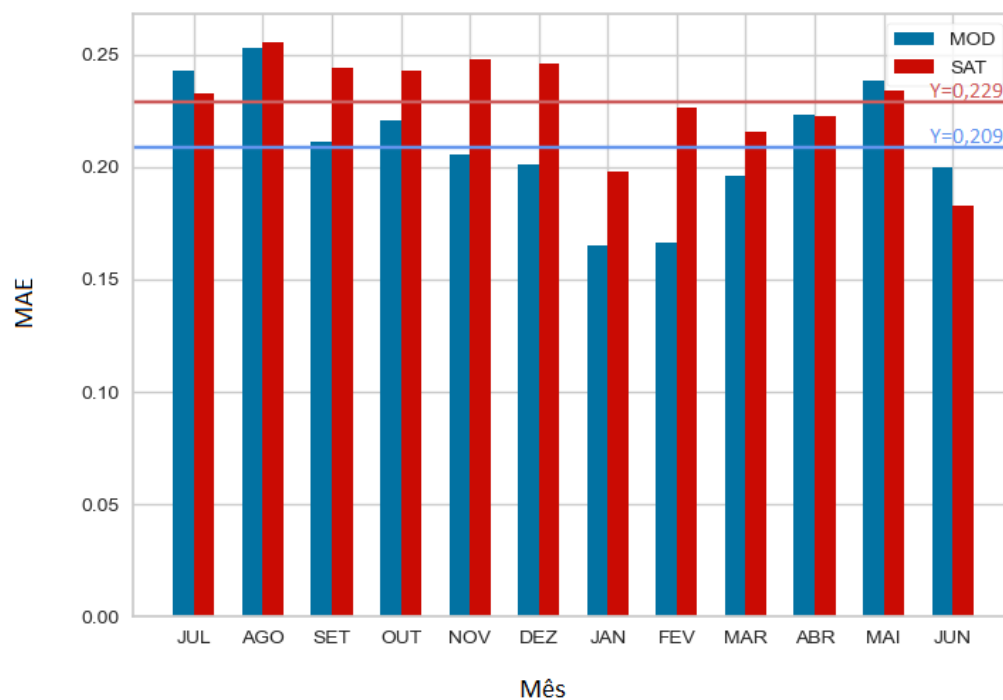


FIGURA C.3: Variação do RMSE por mês de estudo

## Apêndice D - Erros MAE, MBE e RMSE das subáreas

Na tabela D.1 constam os erros de cada subárea definida no Capítulo 1.

	MAE		MBE		RMSE	
	MOD	SAT	MOD	SAT	MOD	SAT
<b>A</b>	0,22	0,22	-0,08	-0,06	0,29	0,30
<b>B</b>	0,19	0,21	-0,06	-0,10	0,25	0,28
<b>C</b>	0,18	0,20	-0,05	-0,13	0,24	0,26
<b>D</b>	0,19	0,20	-0,05	-0,14	0,26	0,26
<b>E</b>	0,19	0,21	-0,07	-0,06	0,26	0,29
<b>F</b>	0,17	0,18	-0,06	-0,12	0,24	0,24
<b>G</b>	0,17	0,18	-0,07	-0,09	0,23	0,24
<b>H</b>	0,18	0,18	-0,06	-0,09	0,25	0,24
<b>I</b>	0,12	0,15	-0,06	-0,10	0,16	0,20
<b>J</b>	0,14	0,17	-0,06	-0,14	0,19	0,22
<b>K</b>	0,21	0,29	-0,09	-0,12	0,26	0,36
<b>L</b>	0,12	0,16	-0,06	-0,11	0,15	0,21
<b>M</b>	0,19	0,27	-0,02	0,08	0,32	0,56
<b>N</b>	0,27	0,31	-0,12	-0,14	0,34	0,39
<b>Média</b>	0,18	0,20	-0,06	-0,09	0,25	0,28